DOCUMENT RESUME

ED 031 706

AL 002 054

By-Rice, Frank A., Ed.
Study of the Role of Second Languages in Asia, Africa, and Latin America.
Center for Applied Linguistics, Washington, D.C.
Pub Date 62
Note-130p.
Available from-Publications Section, Center for Applied Linguistics, 1717 Massachusetts Avenue, N.W., Washington, D.C. 20036 ($2.50).
EDRS Price MF-$0.75 HC Not Available from EDRS.
Descriptors-*Bilingualism, Creoles, Language Planning, *Language Standardization, Language Typology, *Multilingualism, *Second Language Learning, *Second Languages, Sociocultural Patterns, Standard Spoken Usage, Surveys

The purpose of this survey was to investigate the nature and extent of the problem of second language learning as a factor in national development in countries of Asia, Africa, and Latin America. This document is based on the findings of the survey but includes some information not covered in the survey and represents essentially the reaction of half a dozen specialists to some of the problems dealt with in the survey. Essays included are: (1) "Background to Second Language Problems," Charles A. Ferguson; (2) "The Language Factor in National Development," Ferguson; (3) "An Outline of Linguistic Typology for Describing Multilingualism," William A. Stewart; (4) "Mexican and Guatemalan Bilingualism," A. Richard Diebold, Jr.; (5) "Creole Languages in the Caribbean," Stewart; (6) "Lingua Francas, With Special Reference to Africa," William J. Samarin; (7) "Language Situation in East Africa," Ruth E. Sutherlin; (8) "Language Problems in the Rural Development of North India," John J. Gumperz; (9) "Language Standardization," Punya Sloka Ray; and (10) "Sociocultural Change and Communication Problems," Janet Roberts. (DO)

ED0 31706

Edited by Frank A. Rice

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

# Study of the Role of Second Languages

## in Asia, Africa, and Latin America

AL 002 054

Center for Applied Linguistics of the Modern Language Association of America

Washington DC 1962

Edited by Frank A. Rice

# Study of the Role of Second Languages

## in Asia, Africa, and Latin America

Center for Applied Linguistics of the Modern Language Association of America

Washington DC 1962

AL 002 054

## Acknowledgment

This report was prepared pursuant to a contract with the International Cooperation Administration of the United States for its Office of Educational Services.

# CONTENTS

INTRODUCTORY NOTE v

BACKGROUND TO SECOND LANGUAGE PROBLEMS 1
by Charles A. Ferguson

THE LANGUAGE FACTOR IN NATIONAL DEVELOPMENT 8
by Charles A. Ferguson

AN OUTLINE OF LINGUISTIC TYPOLOGY FOR DESCRIBING MULTILINGUALISM 15
by William A. Stewart

MEXICAN AND GUATEMALAN BILINGUALISM 26
by A. Richard Diebold, Jr.

CREOLE LANGUAGES IN THE CARIBBEAN 34
by William A. Stewart

LINGUA FRANCAS, WITH SPECIAL REFERENCE TO AFRICA 54
by William J. Samarin

LANGUAGE SITUATION IN EAST AFRICA 65
by Ruth E. Sutherlin

LANGUAGE PROBLEMS IN THE RURAL DEVELOPMENT OF NORTH INDIA 79
by John J. Gumperz

LANGUAGE STANDARDIZATION 91
by Punya Sloka Ray

SOCIOCULTURAL CHANGE AND COMMUNICATION PROBLEMS 105
by Janet Roberts

## INTRODUCTORY NOTE

This study is one of a series of documents resulting from the Survey of Second Language Teaching conducted by the Center for Applied Linguistics of the Modern Language Association of America during the period December, 1959 - March, 1961 under the provisions of a special grant from the Ford Foundation for this purpose. The aim of the survey was to investigate the nature and extent of the problem of second language learning as a factor in national development in countries of Asia, Africa, and Latin America. It was carried out with the advice of an informally constituted international group of specialists in the field and in active collaboration with the British Council, the French Bureau d'Etude et de Liaison pour L'Enseignement du Français dans le Monde, the Commonwealth Office of Education in Australia, the English Language Education Council, Incorporated, of Japan, and other institutions and individuals in the United States and elsewhere.

One aspect of the survey was the collecting of descriptive data about specific countries and regions. It brought together an extensive body of information about the present position of second or additional languages, especially English and French, as well as about the present extent and status of training in such languages, methods of training, available and potential resources, and current efforts to strengthen those resources. Over sixty country and regional reports were prepared, papers were drafted assessing the activities of resource countries such as Great Britain, France, the United States of America, Canada, and Australia, and a considerable amount of additional information was accumulated in the files. Another aspect of the survey was concerned with bringing together the professional evaluations, opinions, and recommendations of individuals and organizations currently involved in the problem of second language teaching.

The present document is based in the first instance on the findings of the survey but includes some information not covered in the survey and represents essentially the reaction of half a dozen specialists to some of the problems dealt with in the survey. The U.S. Agency for International Development sponsored the preparation of the document because of its concern with some of the basic problems involved in its programs of assistance in the teaching of English as a foreign language in a number of countries.

The opinions presented here have not been submitted to the original international advisory group of the survey, and therefore the authors themselves take full responsibility for the material given in their respective chapters. The Center for Applied Linguistics maintains that all this material, as the considered opinion of specialists, is relevant to fuller understanding of the role of second languages and is worth publication so that it may be available to those interested in it. The Center, in the person of Frank A. Rice, has assumed editorial responsibility for the whole study.

Charles A. Ferguson
Director
Center for Applied Linguistics
April 1962

# Study of the Role of Second Languages

# BACKGROUND TO SECOND LANGUAGE PROBLEMS

by Charles A. Ferguson

To understand the problems involved in the large-scale teaching of languages of wider communication such as English and French in the developing countries it is necessary to become familiar with the very different kinds of language situations which exist in these countries. In a previous study of the Center (Information Series 1 1961) an outline was given of the major language characteristics of Asia, Africa and Latin America. The present study will take up some of the points made there and enlarge upon them, attempting to provide a somewhat more sophisticated theoretical background and much more detailed information on certain sample areas.

Perhaps the most obvious point to deal with in describing the language situation of a country is its linguistic diversity, i.e. the number of different languages spoken and the incidence of multilingualism among the inhabitants. Strangely enough, apart from one article by the anthropological linguist Joseph Greenberg (Greenberg 1956), there has been no attempt on the part of social scientists to devise a numerical index of linguistic diversity which could be used to give a simple characterization of a locality or country, or larger area which could readily be used for comparative purposes. Certainly the planner or administrator in the language education field would welcome an index of this sort which would measure the level of linguistic diversity in one region as compared with that of another.

Greenberg in his article suggested several possible indices of this kind and calculated two of them for parts of Mexico and a few other areas. His simplest index, method A, the "monolingual nonweighted method", measures the probability that two members of the population chosen at random would not speak the same language. The scale runs from 0 to 1, calculated by the formula $A = 1-\Sigma_i(i^2)$ where $\underline{i}$ is the proportion of speakers of each language to the total population. If in a population, to use his example, 1/8 speak language M, 3/8 speak N, and 1/2 speak O then the index figure of linguistic diversity (method A) is .5938. Using this index, a country like Costa Rica where almost 100% of the population speak one language (Spanish), the figure would be very close to zero, while for an area like the Plateau Province of Northern Nigeria, an area noted for its linguistic diversity, the figure is .9539, very close to 1.

This index A does not take into consideration the degree of similarity among the languages in question or the possibility of multilingualism on the part of the inhabitants, and Greenberg suggested several other means of calculating indices (B C D E F G) which take these questions into account. Perhaps his most interesting index, method H, the "index of communication", measures the probability that two members of the population chosen at random would have at least one language in common.

Although Greenberg's indices are promising, they can be of very little immediate help to the planner or administrator, First because accurate information on which to calculate them is not available for most countries and second because they do not take into account the very different importance and range of use of the various languages spoken in a given country. Let us hope, however, that in the next few years more work is done in the development of one or more satisfactory indices of linguistic diversity which could serve as useful tools both for setting a base line in language planning and for measuring achievement of a second language learning policy in a country.

Closely connected with the notion of linguistic diversity is the question of how the various languages in a community are used. To say, for example, that a country has two languages spoken by such-and-such percentages of the population says very little about the language situation, and even a fully quantified specification of this which would reflect the number of speakers of each, the number of lilinguals in the population, the degree of mastery of the second language achieved by various segments of the population, and data on the similarity between the two languages would still leave out the fundamental matter of when and under what circumstances each language is used and what the attitudes of the people are toward the two languages.

For example, France is a country of two languages, where about 97% of the population speak French and about 2% speak Breton; almost all of the Breton speakers are able to speak

French to some extent and many are completely bilingual. The Malagasy Republic also is a country of two languages; about 97% of the people speak Malagasy in one form or another and about 2% of the people speak French; the overwhelming majority of the French speakers also speak Malagasy. In simple quantitative terms these two language situations are similar, but the differences in usage of the respective languages and the attitudes toward them are enormous. In France the minority language is used as a home language in one part of the country while the majority language is used throughout the entire range of national life, including government, education, literature. In the Malagasy Republic, on the other hand, the majority language is the home language of the population and the minority language is used for many facets of national life including most education and many government activities.

A language in a multilingual situation may serve as the identifying language of an ethnic group or it may be the language used only for certain ritualistic purposes in the religion or it may be the means of communication between different speech communities. General studies of multilingualism have attempted to list the variables involved (cf. Lewis 1962) and a number of linguists have commented on the "uses" of language and the existence of "restricted" languages used for limited purposes (cf. Firth 1960) but no satisfactory classification has yet been worked out which can be used to characterize either a language or a language situation from a sociolinguistic point of view. One very tentative proposal was made in a paper on languages and national development (Ferguson 1961) reprinted in this volume and another, somewhat fuller attempt is made in Stewart's paper on typology in this volume. This is a field of considerable importance for the understanding of societies which has been neglected almost completely by sociologists and linguists.

If a typology of languages or language uses were worked out this could be transferred to a typology of nations or other sociopolitical units in language terms. Such a typology was mentioned under the name "national profiles" in the Ferguson paper just cited and is to some extent implicit in the charts by Miss Roberts included in this volume.

The discussion up to this point has been chiefly in terms of the existence of separate languages in a country and the nature and extent of the differences between them. Another question which must be touched upon is the amount of dialect variation within a single language. Some languages are relatively homogeneous, i.e. all speakers talk about the same way, while other languages may show dialect cleavages so great as to offer serious obstances to communication within the speech community. Distinct varieties of speech within a language are called dialects.

It must, of course, be emphasized here that the terms "language" and "dialect" as used by linguists are technical terms without evaluative or emotional character. A dialect in this sense is neither a substandard variety of a language nor a second-class language. A language in this classificatory sense refers to the homogeneous speech form of a whole linguistic community or to a group of such speech forms which are mutually intelligible or constitute a chain of mutual intelligibility. This use of "language", although as basic to linguistic science as the concept "species" is to biology, involved external non-linguistic factors. Mutual intelligibility is difficult and sometimes impossible to test, and it alone does not serve to delimit satisfactorily the units of analysis referred to as languages, since the dialects of some languages are further apart in structure and mutual intelligibility than whole languages are from each other in other cases (Voegelin and Harris 1951, Pierce 1952). It is necessary to include such notions as the existence of a unifying standard form or the attitudes of the speakers toward their language in order to arrive at a fully satisfactory definition of a language. A recent discussion of this question examines the possible kinds of variation and the appropriate definitions, with examples chiefly from South Asia (Ferguson and Gumperz 1960).

It is tacitly assumed by many that one of the features of ideal nationhood is the possession of a standardized national language. The absolute ideal would apparently be a language which has a community of native speakers coterminous with the national boundaries and which has a single accepted norm of pronunciation, spelling, grammar, and vocabulary, used for all levels of speaking and writing, including both a unique national literature and work in modern science.

The actual language situation in a given nation is often far from this ideal, and the selection of a national language and the means of standardizing the language are often among the problems of developing countries. In some nations the selection of a national language has in effect been made by the demographic fact of the country's having one language used by the overwhelming mass of the population (e.g. Thailand - Thai, Somalia - Somali). In some multilingual countries the choice of a national language is simple because of the clearcut dominance of one language in the country (e.g. Iran - Persian, Burma - Burmese, Ethiopia - Amharic). In a few cases (e.g. Indonesia - Malay/Bahasa Indonesia) a minority language has become accepted as the national language. In many countries of Asia and Africa the question of a national language is still troublesome, with no clear solution in sight.

The standardization of a language, i.e. the development of a norm widely accepted throughout the speech community, takes place in various ways. Standardization as a sociolinguistic process is not well understood. There are accounts of the standardization of a number of European languages (cf. Vendryes 1931 pp. 260-279; Meillet 1928) which show quite different sets of historical events in the standardization process of different languages, but there are not enough detailed studies of standardization in various parts of the world and under various circumstances to yield useful generalizations. It is clear that there are limits to what can be legislated or decreed in linguistic change and that certain kinds of change can take place more rapidly than others, but this is far from an adequate understanding of the process. Ray's study in this volume is devoted to a number of aspects of this problem of standardization, especially insofar as it can be a conscious, deliberate operation.

Against these background notions of linguistic diversity, national languages, and language standardization, what is meant by the term "second language" in Asia, Africa, and Latin America? If we assume that it is normal for each human being to learn one language in a "natural" way from his communication partners at an early age (one to five years), this question may be rephrased: When and why does an individual learn another dialect of his language or another language altogether? The answer is twofold. Either the speech community into which he is born uses different dialects or different languages in functionally different ways so that full communication within the community requires a second language, or the individual for one reason or another desires or needs communication with another speech community. In the remainder of this discussion the term "language" will be used to mean either dialect or language.

The existence of marginal, overlapping, or anomalous cases does not invalidate this general formulation: an individual normally learns the primary language of his speech community at an early age and adds to it (a) additional languages required for full communication within the community and (b) additional languages needed for communication outside the community. Insofar as additional languages learned as (a) or (b) are regional or national languages or are required for access to modern science and technology or are needed for international communication they are directly relevant to the social, economic, and educational development of the nation.

Second language learning takes place either by relatively informal, unplanned imitation and use in actual communication situations or by formal study in a system of education. The impression of specialists in the language field is that languages learned by the informal "using" method are learned faster, more completely, and with greater retention than languages learned as subjects in school or special educational situations.

If this is so, then experimentation should be undertaken to find out the basic facts here. The problem for the educator is how to make the acquisition of languages through formal education - which must be the chief agency - either as much like the more natural learning as possible or else to discover and use methods of language learning different from the natural ones but superior in results.

## REFERENCES

Ferguson, C. A. 1962. "The Language Problem in National Development" Anthropological Linguistics 4:1.23-28.

Ferguson, C. A. and Gumperz, J. J. 1960. "Introduction" in Linguistic Diversity in South Asia. Bloomington: Indiana University. Pp. 1-18.

Firth, J. R. 1960. "The Teaching of English Overseas in Relation to Uses" [London] Multilithed. 7 pp.

Greenberg, J. H. 1956. "The Measurement of Linguistic Diversity" Language 32.109-115.

Information Series 1 1961. Second Language Learning as a Factor in National Development in Asia, Africa and Latin America. Washington: Center for Applied Linguistics. v, 18 pp.

Lewis, E. G. 1962. "Conditions Affecting the 'Reception' of an 'Official' (second/foreign) Language". London: CCTA/ CSA. Mimeog. 27 pp.

Meillet, A. 1928. Les langues dans l'Europe nouvelle. 2nd ed. Paris.

Pierce, J. E. 1952. "Dialect Distance Testing". International Journal of American Linguistics 18.203-10.

Vendryes, J. 1931. Language: a Linguistic Introduction to History. Tr. P. Radin. London.

Voegelin, C. F. and Harris, Z. S. 1951. "Methods for Determining Intelligibility Among Dialects of National Languages" Proceedings of the American Philosophical Society 95.322-9.

# THE LANGUAGE FACTOR IN NATIONAL DEVELOPMENT

by Charles A. Ferguson

0. Introduction
1. Development
2. Nations
3. National profiles

0. Social scientists of various disciplines are concerned with the concept "national development", in particular, of course, economists and political scientists, but to a lesser extent scholars in other fields.[1] Structural linguistics, however, in spite of its concern with diachronic matters, has been resolutely opposed to any developmental or evolutionary approach in linguistic analysis. The purpose of this paper is on the one hand to suggest the relevance of "national development" for linguistic analysis and on the other hand to point to linguistic aspects of national development as it is studied by social scientists in other fields. The approach followed here has resulted from the work of the Survey of Second Language Learning in Asia, Africa, and Latin America which was carried out by the Center for Applied Linguistics in collaboration with outside specialists from the United States and other countries.[2]

1. Of the many scales which could be developed for measuring language "development" in a way which might correlate usefully with nonlinguistic measures of development two seem particularly promising: the degree of use of written language and the nature and extent of standardization. Scales suggested here represent a modification of the viewpoint of Heinz Kloss.[3]

---

From Anthropological Linguistics 4:1.23-27 (1962). Reprinted by permission of the author and the editor of Anthropological Linguistics.

1.1. The 3000 or more languages currently spoken vary in the use of a written form of the language from cases in which the language has never been written to languages with an enormous and very varied use of written forms. It is difficult to arrange these cases in a simple, linear progression, partly because of the complex variation and partly because of the great range in the amount of use. As a first approximation to a useful scale, we suggest the scheme

W0. - not used for normal written purposes
W1. - used for normal written purposes
W2. - original research in physical sciences regularly published

A convenient set of criteria for establishing "normal" use of the written language is as follows. (a) The language is used for ordinary interpersonal epistolary purposes. People write letters in it. (b) The language is used in popular periodicals. Newspapers appear in it. (c) The language is used in books not translated from other languages. People write and publish books in it.

Languages with rating "0." include languages such as Modern Aramaic for which no orthography has been suggested and which has no representative writing by members of the speech community, as well as languages like Tuareg where use of the writing system is limited to special and marginal purposes, or Lugbara where an orthography has been suggested which has been used in some dictionaries, grammars, textbooks, Bible translations and the like, but which has not yet become widely used in the community. Most languages are still at this level although there is a steady stream of languages moving up to level "1".

Examples of languages at level "1" include Amharic, Thai, Slovenian. Many languages in category "1" have a substantial publication output. Languages with relatively small output often have a considerable amount of poetry, folkloric material and translations. Only a few languages fall in category "2". These are often also languages widely used for intercommunication by other speech communities. The number of languages belonging in this category, however, is steadily increasing and this suggests the need for a possible additional level.

W3. - languages in which translations and resumés of scientific work in other languages are regularly published.

1.2. The establishment of a scale of standardization is much more difficult because there are at least two dimensions involved, one of which is itself quite complicated. One such dimension is the degree of difference between the standard form or forms of a language and all other varieties of it. This difference may be very small or very great independently of the other dimension, which is the nature of the standardization and the degree to which a standard form is accepted as such throughout the community. The simplest approach seems to be to set up the end points of the scale as St$^0$. and St$^2$. Zero refers to a language in which there is no important amount of standardization. As an example, we may take Kurdish, where there is a considerable dialect variation but where no form or forms of the language has received wide acceptance as a norm among people who do not speak it themselves.[4] At the other end of the scale we have what may be regarded as the "ideal" standardization. The term "ideal" is not inappropriate because individuals concerned with the development of their nation who make proposals for change in the language situation generally seem to make proposals aimed at achieving this "ideal" standardization even though they rarely state the desired goal explicitly. Category "2." refers to a language which has a single, widely accepted norm which is felt to be appropriate with only minor modifications or variations for all purposes for which the language is used. Differences between regional variants, social levels, speaking and writing, and so on, are quite small. An example of category "2."is Swedish, where the difference between the written and spoken standard is appreciable, but relatively minor and growing less, and where none of the original dialects are too far removed from the standard.

Category "1."requires considerable sub-classification to be of any use. Whatever scheme of classification is developed, it will have to take account in the first instance of whether the standardization is unimodal or bi- or multimodal, and in the case of more than one norm, the nature of the norms must be treated. Armenian may serve as an example of a bimodal standardization where the standards are essentially regional, East Armenian and West Armenian, both being used for normal written purposes. Greek may be cited as an example of a bimodal standarization based on a "vertical" or role differentiation, one being used for ordinary conversation and the other for most written and formal spoken purposes. Serbo-Croatian has two norms based to a large extent on religio-cultural differences. Norwegian is an example of a bimodal standardization based on neither of these. As a single example of a more complicated case of standardization, we may mention the whole Hindi-Urdu complex with its regional standards in addition to the religio-cultural split.[5]

2. Linguists have generally operated with the concept of speech community "a group of people who use the same system of speech-signals"[6] as the locus of linguistic behavior, although recently some attempts have been made to deal with multilingual communities.[7] Only rarely has the concept nation been utilized by linguists for this purpose, and then generally for describing certain features of language used in Europe.[8] From many points of view, however, it is desirable to use the nation as the basis for general sociolinguistic descriptions: communication networks, educational systems, and language "planning" are generally on a national basis and national boundaries play at least as important a role in the delimitation of linguistic areas as any other single social barrier. In the description of the language situation of a given nation two fundamental points must be treated, the number of languages and the relative dominance of languages.

2.1. In determining for taxonomic purposes the number of languages spoken in a country it seems advisable to distinguish between major and minor languages. A definition of major language which has proved useful in the work of the Survey is: a major language of a nation is a language spoken by at least ten million people or one-tenth of the population. The number of major languages in a nation may vary from one to a dozen or more. It seems likely, however, that the important categories are: one major language (e.g. Thailand, Costa Rica, Holland), two major languages (e.g. Canada, Belgium, Paraguay), and three or more major languages (e.g. Switzerland, Nigeria, India).

2.2. In a nation with more than one major language, it is often true that one is clearly dominant over the others or, in some cases, several languages are dominant over the others. One indication of dominance is numerical superiority: one language is dominant over others if it is spoken by more than half the population of the country. Another important indicator of dominance is the extent to which a given language is learned by native speakers of other languages in the country. For example, Persian and Pashto are spoken by about the same number of people in Afghanistan, but Persian is often learned as a second language by speakers of Pashto and other languages in the country, while Pashto, in spite of official government support and formal classes, is rarely learned well by speakers of other languages in the country. A third indicator of language dominance is the use of one of the languages of the nation for such clearly national uses as publication of official texts of laws or decrees, medium of instruction in government schools, normal channel of military communication.

Full agreement among these three indicators provides the "normal" form of national language dominance. Cases where these indicators are not in agreement seem generally to have serious social tensions connected with language problems.

2.3. In many nations, especially in Asia and Africa, languages of wider communication (LWC) such as English and French, play an important role in the national language situation and this must be separately assessed. For one thing, the LWC may be the language used for the clearly national purposes listed under the third indicator of dominance. In addition it may be an LWC rather than one of the local languages which is used as the means of access to scientific and technological knowledge or to communicate with other nations in the expanding network of international communication.

3. With the fairly simple machinery outlined in the previous two sections, it is possible to draw up for a nation a profile which will be of value for comparison with other non-linguistic indices of development. The great drawback here is not the theoretical complexity or even the practical man-hours of work involved. It is the lack of reliable data for most nations.

The preparation of a national sociolinguistic profile in the sense described here calls for putting down on paper about a given nation the following information: how many major languages are spoken; what is the pattern of language dominance; are there national uses of a LWC; for each major language spoken in the country, what is the extent of written uses of the language (W0. - W2.) and what is the extent of standardization (St0. - St2.) and its nature (multimodal? range of variation from the norm?).

Even with the preliminary profiles which can be put on paper on the basis of currently available data, it is evident first that there is a wide range of variability although certain types are quite common, especially the one nation - one major dominant language W1. St2. It is also evident that certain types of profile occur only in underdeveloped countries, especially the one with no dominant language and an LWC used for national purposes as well as for access to science and international communication. Before any useful theorizing can be done, it is necessary to collect the data and prepare reliable national profiles. The possibility of significant conclusions arising from such study seems very promising.

NOTES

1. As examples of recent works concerned with the theory of national development we may cite W. W. Rostow, The Process of Economic Growth (New York, 1952); G. A. Almond and J. S. Colemen (eds.), The Politics of the Developing Areas (Princeton, 1960). One work which emphasizes the role of communication in the concept of nationhood is K. W. Deutsch, Nationalism and Social Communication (New York, 1953).

2. For an account of the Survey and some of its results, see Second Language Learning...in Asia, Africa, and Latin America (Washington: Center for Applied Linguistics, 1961).

3. Cf. H. Kloss, Die Entwicklung neuer germanischer Kultursprachen (Munich, 1952), in particular pp. 24-31,"Stufenfolge des Ausbaus eines Idioms zur Kultursprache."

4. Kurdish may not be a fully satisfactory example since the dialect of Suleimaniya is beginning to be accepted as a norm by a considerable segment of the Kurdish speech community. Cf. E. N. McCarus, A Kurdish Grammar (New York), 1958) p. 1; D. N. Mackenzie, Kurdish Dialect Studies I (London, 1961) p. xviii.

5. Cf. J. Gumperz and C. M. Naim, "Formal and Informal Standards in the Hindi Regional Language Area," in C. A. Ferguson and J. Gumperz (eds.), Linguistic Diversity in South Asia (Indiana University, RCPAFL 13, 1960)

6. L. Bloomfield, Language (New York, 1933) p. 29.

7. Cf. U. Weinreich, Languages in Contact (New York, 1953) pp. 83-110.

8. See, for example: L. Dominian, The Frontiers of Language and Nationality in Europe (New York, 1917); A. Meillet, Les langues dans l'Europe nouvelle (Paris, 1928); S. Rundle, Language as a Social and Political Factor in Europe (London, 1946).

# AN OUTLINE OF LINGUISTIC TYPOLOGY FOR DESCRIBING MULTILINGUALISM

by William A. Stewart

Of the special social and technical problems which beset new and developing nations in particular, a sizable number turn out to be directly related to language in some way. Such linguistic problems include those of widespread illiteracy, the lack of a standardized national language, the need for pedagogical tools in locally adapted language teaching, and the lack of modern technical vocabularies in languages which must suddenly be employed for communicating technical and scientific knowledge. Linguistic problems such as these can and do occur in areas with linguistic uniformity and diversity alike but it is especially in the latter case that they may be characterized by greater complexity, with their solutions rendered especially difficult.

National multilingualism, that is, the use of more than one language within the same national territory, exists to some extent in all the major areas of the world. It is especially in Asia, Africa, and Latin America, however, where--often as the linguistic aspect of a still vigorous ethnic or cultural pluralism--multilingualism has given rise to national communication problems of a serious enough nature to have prompted a number of national governments to initiate remedial programs. Although these programs have varied greatly in their details, they have usually been based upon one of two fundamentally different policies:

1. The eventual elimination, by education and decree, of all but one language, which remains to serve for both official and general purposes.

2. The recognition and preservation of important languages within the national scene, supplemented by universal use of one or more languages to serve for official purposes and for communication across language boundaries.

The first of these approaches aims at de-emphasizing linguistic diversity as much as possible, and is usually part of a more general policy of total assimilation to the "national" culture of all ethnic and cultural minorities. The second approach is more tolerant of multilingualism, and is merely concerned with bringing about the additional use of certain languages for inter-communication and for national representation. This usually parallels a policy of officially recognizing cultural pluralism as a fundamental part of the nation's makeup. The extent to which either policy is a more suitable one depends largely upon the nature of the particular multilingual situation to be dealt with. Generally speaking, the discouragement of linguistic pluralism is most feasible in countries where essentially one language has already come to enjoy a high degree of usage and prestige, while the acceptance of linguistic pluralism is a more realistic course for nations which are subdivided into a number of established and to some extent culturally autonomous linguistic communities.

Almost every administrative, technical, or pedagogical program which has been undertaken in order to deal with multilingualism has met with problems in the form of unpredicted reactions either for or against changes in the linguistic status quo. For example, there have been cases where the speakers of one language have resisted being taught some other one because of a feeling that the second language, even though it might have had a far larger number of actual speakers, would be somehow "less appropriate" for general use. Or, there have been other cases where individuals have preferred to devote their efforts toward attaining fluency in some language other than their native one--in spite of the greater difficulties involved--merely because the second language was regarded by them as being "more worthy" than their native language. Where social reactions of this type have not been foreseen, it has been largely because not enough information was sought in advance about the different ways in which the various languages in a multilingual situation may intermesh with social attitudes and behavior patterns. Because of this, both scholars and policy makers in the language field would benefit from a familiarity with at least the more recurrent sociolinguistic phenomena which may characterize multilingualism, especially those differences in the ways in which individual languages are ranked and used by people. Particularly useful insights into the sociolinguistic constants and variables of multilingualism may be obtained through the description and comparison of different national case histories. Indeed, governmental administrators might find a comparative context a useful one in which to evaluate variations in the progress or outcome of past language-oriented programs as well as in the planning of new ones.

In the present Outline, a sociolinguistic typology has been developed especially for use in the description of national language situations. It is designed to indicate in somewhat abbreviated form the main kinds of functional and distributional relationships which different languages may have in the national linguistic scene.[1]

## A. Language Types in Multilingualism

One important factor in the typology of multilingualism is an essentially attitudinal one, which will be referred to in this Outline as language type. It derives from the apparently general phenomenon that, owing to attitudes which people have come to adopt toward certain socio-historical attributes of language, different languages may be accorded different degrees of relative social status. Because of this, languages which occur together may not stand in a relationship of social equality with respect to one another. Language type in this sense is often a major factor in determining whether or not a language will be accepted by the members of a national society as suitable for some specific rôle, such as for use in education or as an official language. The specification of the sociolinguistic type to which a particular language belongs requires taking into account several different facts about it, since no single social or historical attribute seems to be wholly decisive in determining a language's potential social status. Languages may be differentiated into types in terms of four attributes which demonstrate a consistent tendency to affect social attitudes toward them. These are:

I. Historicity, i.e. whether or not the language is the result of a process of development through use.

What makes a language obviously historical is its association with some national or ethnic tradition.

II. Standardization, i.e. whether or not there exists for the language a codified set of grammatical and lexical norms which are formally accepted and learned by the language's users.

Formalized grammars and lexicons tend to impart a generally accepted status of "legitimacy" to languages which possess them.

III. Vitality, i.e. whether or not the language has an existing community of native speakers.

Other things being equal, the recognition that a language has native speakers will have the effect of increasing its social importance for other people.

IV. Homogenicity, i.e. whether or not the language's basic lexicon and basic grammatical structure both derive from the same pre-stages of the language.

Most languages of the world are homogenetic, but there are languages of which the basic grammar comes from one source and the basic lexicon from another.

The way in which these socio-historical attributes combine to produce viable and extant language types is shown in the following chart. A plus sign indicates the presence of one of the variables, and a minus its absence. The resultant language types appear in descending order of potential social prestige.

| Attributes | | | | Language Type | Type Symbol |
|---|---|---|---|---|---|
| I | II | III | IV | | |
| + | + | + | +<br>- | Standard | S |
| + | + | - | + | Classical | C |
| + | - | + | + | Vernacular | V |
| + | - | + | - | Creole | K |
| + | - | - | - | Pidgin | P |
| - | + | - | +<br>- | Artificial | A |
| - | - | - | +<br>- | Marginal | M |

*Standard* (S) and *classical* (C) are the two language types with the highest universal prestige. Their codified norms--reinforced by literary models in many cases--tend to encourage a relatively high degree of uniformity in usage, even where the languages' users are geographically or socially separated. In some cases, allowances for certain kinds of regional variation may even be made for within the formal grammatical and lexical conventions. These two language types are the ones which are the most likely to be adopted as official languages in newly formed nations. An example of a *standard* type language would be English as it is spoken or written by most educated users of the language.[2] Perhaps the best example of a *classical* type now in use is that of literary Arabic. Outside of Malta, which has an Arabic *standard* (i.e. Maltese), it is the only standardized form of Arabic, and exists in each country beside one or more related vernaculars. Greece provides an interesting example of a *standard* (i.e. Dhimotiki) and a *classical* (i.e. Katharevusa) used side by side. It is possible for a *classical* to become or give rise to a modern *standard* language by regaining a community of native speakers, cf. Israeli Hebrew, a standard developed from the *classical* language.

*Vernacular* (V) languages are usually learned as first, or native languages, although in some cases they may be informally acquired as second languages by individuals. Due to their lack of formalized grammars and lexicons (at least ones accepted as authoritative by the languages' users[3]), *vernacular* languages are almost always ranked lower in prestige than *standard* or *classical* languages. The fact that they often have large numbers of monolingual speakers does tend to give them some social or political weight. Most tribal languages of Africa and the Americas are *vernacular* types. So also are many speech forms related to *standard* languages but which, though having no formalized grammars of their own, are functionally and structurally autonomous enough to rank as separate languages, cf. Yorkshire dialect or Egyptian colloquial Arabic. Because of their lack of standardization, *vernacular* languages are generally not used for formal writing. Of course, a *vernacular* may eventually become a *standard* through the gradual codification and acceptance of its grammar and lexicon.

*Creole* (K) and *pidgin* (P) languages are both the result of the development of a secondary language for wider communication in certain kinds of social and linguistic contact situations where grammatical and lexical material from different sources became fused. A pidgin is such a language in its primary stage, when it is spoken only as a second language. At times such a language may acquire native speakers and undergo further development, and so become a creole. One widely used pidgin which is fast becoming a creole is Neo-Melanesian, or Melanesian Pidgin English. Several closely related French-based creole languages are spoken in the Caribbean, and English-based creoles are spoken in Surinam (i.e. Sranan and Saramakkan) and in Sierre Leone (i.e. Krio). Like vernaculars, creoles may eventually become standard languages.

Artificial (A) and marginal (M) languages are alike in that they are both ad hoc creations. The former usually have a highly rigidified and simplified structure which has been contrived by their individual authors. Marginal speech forms have much less of an established structure, being simply tools of private communication between very limited numbers of individuals who are frequently in contact, but who know no language in common. Examples of a few artificial languages which have become well known are Volapük, Esperanto and, more recently, Interlingua. Marginal languages have been quite common in colonial Asia and Africa as "household languages" developed between Europeans and their servants.[4]

### B. Language Functions in Multilingualism

Another factor of importance in the typology of multilingualism is that of the function of each language as some kind of medium of communication within the nation. Different languages, even when of the same language type, may have differing functions; some may occupy highly specialized rôles, while others may be used for very generalized purposes. Although differentiation between linguistic functions can be made as refined as the descriptive goals warrant, the following functional categories have been considered adequate enough for the purposes of general description and comparison:

Official (symbol: o)[5]; the use of a language as the legally appropriate one for all politically and culturally representative purposes. In many cases, the official function of a language is specified constitutionally. In addition to its representative rôle, an official language is almost always among those used for educational and literary purposes, and for wider communication within the nation.

*Group* (symbol: g); the use of a language primarily by the members of a single ethnic or cultural group or sub-group. So strong can be the association between language and group, that at times languages with a *group* function may often serve as informal criteria for ascertaining group membership. The fact that a language is associated in this way with a particular social group may tend to create resistence on the part of out-groupers to the adoption of that language for more general purposes--particularly if the group with which the language is associated is otherwise strongly segmented from the rest of the national society.

*Wider communication* (symbol: w); the use of a language, other than an official one, for communication across language boundaries for purposes of trade and commerce within the nation.

*Educational* (symbol: e); the use of a language, other than an official one, as a medium of instruction at some level of the educational system.

*Literary* (symbol: l); the use of a language, other than an official one, primarily for literary or scholarly activities.

*Religious* (symbol: r); the use of a language primarily in connection with the practice of a religion.

*Technical* (symbol: t); the use of a language primarily as an access to international technical and scientific literature.

Under various conditions, the same language may occupy more than one functional slot. Although the correlation between language type and language function is flexible enough so that it must be stated for each specific case, it is generally the case that the *official* (o), *literary* (l), and *religious* (r) functions will almost always be occupied by either *standard* or *classical* languages, the *technical* (t) function by *standard* or *artificial* languages, and the *group* (g) function by *standards*, *vernaculars*, or *creoles*.

Multilingual situations tend to stabilize so that different languages used together are often functionally non-competitive. This is accomplished by any two languages being in complementary distribution in some social or geographical sense. For example, two languages with a *group* function or with a *religious* function will by definition virtually always be used by different social groups or for different religions, respectively. On the other hand, two languages with different functions, for instance, one with a group and the other with a religious function, may be used by the same people but under different circumstances; the language with the group function will be used for general purposes, excluding religious ritual, while the religious language will be used only for the latter purpose.

Problems of linguistic conflict occur when this reciprocal network is broken down, either by a natural historical process, as in the development of simultaneous pressure from two linguistically different cultural sources, or by direct administrative intervention, such as when an attempt is made to replace a vernacular language by a different standard language in a particular group function, or by political competition between two different languages for a wider communication or official function in a given area. In such cases, either one of the languages will eventually be eliminated from use, even though the remaining one may be heavily influenced by the other in structure and lexicon.

## C. Diglossia

Because of differences in language types and language functions which can occur in any case of national multilingualism, linguistic pluralism can exhibit a network of very complex relationships. The situation is further complicated by the fact that there are two very different kinds of associative relationships which two languages may stand in with respect to each other. So far, it has been assumed that the normal relationship between any two languages was: (a) general mutual unintelligibility, and (b) no marked mutual identification of the two languages. This may be considered the "normal" bilingual relationship.

In general, a historical relationship between two languages does not necessarily guarantee either any appreciable degree of mutual intelligibility or mutual identification, especially if the relationship is one of the remote past.[6] Where, however, two languages are historically related closely enough so that there is an obvious correlation between them at some structural level--usually lexicon--of both, and where the languages are in complementary functions, then they may come to be regarded as variants, in some way, of each other. This special kind of language relationship is usually termed diglossia in sociolinguistics.[7] In some cases of diglossia, two languages may be in such a subtle reciprocal function that they may even serve as different style levels in the same discourse situations.[8]

In most cases of diglossia, the two languages represent different language types as well, and the resultant hierarchical relationship tends to intensify their reciprocity and to inhibit the drift of one language into the functional slot of the other.

A diglossia relationship between languages can be symbolized by the notation "Language *A* ~ Language *B*", with the language of higher prestige always listed first.

A sample of this and other symbols suggested in this *Outline* can be furnished in the form of a notational résumé of the linguistic situation in the State of Louisiana:

| | | |
|---|---|---|
| English | (So) | |
| French | (Sl) | (French ~ Acadian, Negro Creole) |
| Acadian | (Vg) | (French ~ Acadian) |
| Negro Creole | (Kg) | (French ~ Negro Creole) |
| Latin | (Cr) | |

This can be interpreted: "there is a standard language, English, which is the official language of the area. There is also another standard language, French, which is used primarily for literary and scholarly purposes. In addition, there are two languages used by special social groups: a vernacular called Acadian and a creole called Negro Creole. Both stand in a diglossia relationship to standard French, which serves as the language of higher prestige. Lastly, there exists a classical, Latin, which is used primarily in carrying out the ritual of a particular religion."

## NOTES

1. Most linguistic typologies have been concerned with indicating historical relationships between various languages. However, historical relationship is not necessarily significant as such in multilingualism, particularly if the related languages have evolved sufficiently along different lines. Thus English, Bengali, and Hindi are all related languages,

but this fact appears to have no effect on the problems involved in their coexistence in multilingual India.

2. In a strict sense, standardization is independent of writing, since there have apparently been cases of unwritten languages which have nevertheless been standardized on the basis of traditionally recited and preserved oral models. Also, any language can be technically reduced to writing for certain specialized purposes, e.g. linguistic description, the publication of biblical texts, poetry, etc., without any of these writing techniques becoming established among the language's real users. However, in view of the fact that writing is so frequently a characteristic of present-day standard languages, it will be generally included in the definition of standardization here, but only in the sense of "an established writing system, together with an appropriate set of orthographical conventions, which has become part of the language's codified grammar." The reader who knows German will find a useful discussion of the relationship of writing to standardization in Chapter I of Heinz Kloss, _Die Entwicklung neuer germanischer Kultursprachen von 1800 bis 1950_, Munich, 1952.

3. Linguists hold that all languages have a grammatical structure and a vocabulary, and that these are subject to scientific analysis and description. The question is merely whether or not the culture with which the language is traditionally associated has developed a formally accepted set of rules about the way in which the language is supposed to behave.

4. This characterization of _marginal_ languages is essentially an expansion of the one in J. Berry's typology of Sierra Leone multilingualism in his "Pidgins and Creoles in Africa," _Second Meeting of the Inter-African Committee on Linguistics; Symposium on Multilingualism_, Brazzaville, 16-23 July, 1962 [advance copy, p. 2]. The process involved in the formation of _marginal_ languages may possibly be akin to that of some _pidgins_, but of course on a smaller scale. For this reason, _marginal_ speech forms deserve additional study under controlled conditions.

5. Upper case symbols have been used for coding language types, and lower case symbols for language functions.

6. Even where two languages are fairly closely related

historically, their identification with different and sometimes rival cultures may cause attitudinal interference with any mutual intelligibility which might be feasible because of similarities in their structures. This phenomenon is discussed in Hans Wolff, "Intelligibility and Inter-ethnic Attitudes" in Symposium on Urbanization and Standard Language (= Anthropological Linguistics, Vol. 2, No. 3, March 1959).

7. For a general discussion of diglossia, with examples from four similar but unrelated cases, see Charles A. Ferguson, "Diglossia," in Word, vol. XV, No. 2 (August 1959).

8. For examples of this particular aspect of diglossia taken from two different cases, see Haim Blanc, "Stylistic Variation in Spoken Arabic: a Sample of Interdialectal Educated Conversation" (in Charles A. Ferguson, ed., Contributions to Arabic Linguistics, Cambridge, Harvard University Press, 1960), and William A. Stewart, "The Functional Distribution of Creole and French in Haiti" (to appear in Georgetown University Monograph Series on Languages and Linguistics, No. 15, Thirteenth Annual Roundtable of Languages and Linguistics.)

# MEXICAN AND GUATEMALAN BILINGUALISM

by A. Richard Diebold, Jr.

## 1. INTRODUCTION

From the late fifteenth through the seventeenth centuries, a technologically superior Spanish task force busied itself with the conquest and colonization of much of the American Indian territory which makes up modern Latin America. By the middle of the sixteenth century, it had achieved hegemony over the various Indian groups in most of that area, including the region which comprises the modern republics of Mexico and Guatemala, a region which I will hereafter refer to as Middle America.

Linguistically, Spanish hegemony created in Middle America a situation of multiple language contact in which Spanish soon assumed the role of a dominant lingua franca (Diebold, 1962), and Middle American Indian languages today constitute a grouping of languages on the basis of their common tradition of contact with and linguistic interference from a single dominant language.

But it is upon the social role played by the Spanish lingua franca that I want to base my remarks here. It is surprising how little study has been directed to this question (Diebold, 1958). In ways which should become clear below, culture change in rural Middle America is very much bound up with becoming bilingual in Spanish, which is the national lingua franca. It is my hope that the following discussion of Middle American bilingualism will offer to those persons engaged in the teaching of lingua francas (such as English and French) as secondary languages, an insight into the relationship between such a secondary lingua franca and the primary languages with which it is in often competitive contact.

The following section will describe some of the salient features of Middle American bilingualism. A final section will discuss some of the "grassroots" details of bilingualism, drawing upon a study of bilingualism which I undertook in the Huave Indian community of San Mateo del Mar, in the Mexican state of Oaxaca (see Diebold 1961a, 1961b).

## 2. SOME FEATURES OF MIDDLE AMERICAN BILINGUALISM

After the initial period of conquest, Spanish colonization was marked by a concern with dividing up the Indian population for purposes of administration and taxation. There thus emerged very early an interest in the taking of censuses, and the communications problems inherent in such a linguistically diverse area early forced attention on the census quantum of the language which an individual spoke. The interest in collecting data on primary language and bilingual skills has persisted through until the latest national censuses taken in Mexico and Guatemala.

There has been controversy over the sociological classification of the Middle American population into Indian and Mestizo elements (e.g. Caso 1948; Gamio 1948; Gillin 1949; Goubaud Carrera 1946, Luís Arriola 1950; Nash 1957; Redfield and Tax 1952); and one of principal diagnostics used in this classification is the language spoken. An appreciable segment of the population in certain areas learns Spanish only as a secondary language, or not at all; the primary language is an American Indian language. Often enough, the sociological classification seizes as its criterion the distinction that the "Indian" first acquires an American Indian Language in childhood. It thus becomes an easy task not only to decide who is and who is not an Indian, but to proceed to finer distinctions according to the particular language spoken; and thus one speaks of Mayas, Zapotecs, Mixtecs, and so on. The sociological term is in part a linguistic index.

It has been asserted that an individual can move upward socially from an Indian to a Mestizo group by a mere juggling of language symbols, i.e. the chief item of his Indian heritage which the individual in transition must consider is his language: he has to learn Spanish. Many previously isolated communities in Middle America which are monolingually Indian-speaking find themselves engaged in greater social contact with and participation in the national life, and an inevitable concomitant is a greater degree of bilingualism and even shifting to Spanish (Aguirre Beltrán 1957; Gamio 1948; Gillin 1949; Goubaud Carrera 1946c). With bilingualism (the learning of Spanish as a secondary language) and acculturation (to the Mestizo model) the individuals affected often leave their Indian community to join the Mestizo society. This accounts in part for the conservatism and uniformity manifested by many of these communities: they lose their culture marginals (and hence many of their bilinguals) to the ever-growing national society.

But for those communities whose members are predominantly monolingual in an Indian language, there looms a communications problem, for any venture in bicultural behavior involves the acquisition of the language of the dominant culture.

Greenberg (1956:110) has summarized the problem as follows: That it is expected "that areas of high linguistic diversity will be those in which communication is poor, and that the increase in communication that goes with greater economic productivity and more extensive political organization will lead typically to the spread of a lingua franca, whether indigenous or imported, resulting in widespread bilingualism and the ultimate disappearance of all except a single dominant language."

Additional factors are relevant: First, there are no Spanish Indian creoles or pidgins reported from Middle America. Second, although bilingualism today usually involves a primary Indian language and the acquisition of Spanish as a secondary language, there is still a sizeable incidence of Indian-Indian bilingualism, such as Otomí-Nahuatl.

## 3. SOME DETAILS FROM A BILINGUAL COMMUNITY

The 1950 Mexican national census showed that the municipio of San Mateo, a community of speakers of Huave, had a population of 3611. Unofficial calculations placed the percentages of bilinguals in the total population at "less than" 20%.

One of the primary field-problems was to retake a village-wide census. This resulted in a household census which sought information on household constituency and the kinship relationships involved, and the age, sex, place of birth, and language skills of the various individual members of the household. Moreover, the census inquired about the household head's contacts with the outside world and about his occupation.

In this census, a person's language skills were categorized by highly impressionistic measurements. Native Spanish-speakers assisted in taking the census. They were asked to classify as coordinate bilinguals any Mareños (i.e. inhabitants of San Mateo) who appeared to speak Spanish natively or with little accent. Mareños whose Spanish was productive but which exhibited noticeable interference from Huave were to be classified as subordinate bilinguals. Persons unable to make complete meaningful utterances in Spanish were classified as monolinguals. The breakdown of the total of 2934 speakers into these three categories was as follows: 175 coordinate bilinguals (6% of the total), 392 subordinate bilinguals (13%), 2369 monolinguals (81%). Certain problems arose in the classification of a large number of monolinguals who, in their minimal use of Spanish, seemed really to be incipient bilinguals; this problem is treated in Diebold 1961b.

If we examine the data of the village-wide census, three significant social correlates to San Mateo bilingualism emerge:
(1) By sex. Bilingualism is predominantly a male skill. Of the 175 individuals classified as coordinate bilinguals, 146 (or 86%) were males; of the 392 classified as subordinate bilinguals, 273 (or 70%) were males. Of the 49 female coordinate bilinguals, 38 were young adults or adolescents who resided in households the head of which (and often the mother) was bilingual. Of the 119 female subordinate bilinguals, 91 resided in households in which the head was bilingual.

(2) By age. Bilingualism is predominantly an adult or adolescent skill. No one interviewed reported that he had acquired Spanish as a primary language or had learned it simultaneously with Huave. This is borne out consistently by the census data: of the 175 coordinate bilinguals listed, 121 were age 18 or older; of the 392 subordinate bilinguals, 264 were age 18 or older. The bilinguals under age 18 (34 coordinate and 128 subordinates) are all from households having at least one bilingual adult. Of the 34 coordinate bilinguals under age 18, 23 had bilingual mothers.

Factors of sex and age taken together indicate that the majority of bilinguals will not pass on this linguistic heritage to their children.

(3) By occupation. Bilinguals appear in comparable numbers from the three male occupation categories: farmer, fisherman, or artisan. The direct correlation is not to the occupation but rather to the degree to which the occupation obliges the individual to establish contact with outside Spanish-speaking communities. While frequency of contact did not correlate for the monolingual group, it did for both the bilingual groups. Two variables in one's economic life seem to increase the intensity of the contact: (a) the individual's having a salable surplus, and (b) the recurrent need for particular crucial goods from the larger outside markets.

In summary, we can state that Spanish is rarely the language of the home and that it is always secondary. The situation with respect to the age of acquisition and the sex of the individual is in fairly stable equilibrium: the Huave child cannot often acquire Spanish as a childhood language since few mothers would themselves be bilingual. Contact with the Spanish-speaking outside world seems to be the crucial factor involved. Spanish is learned late in adolescence or in adulthood, and chiefly outside the community. Individual Huaves acquire various degrees of proficiency in Spanish outside their home community; upon return to the home, they have little or no opportunity to speak Spanish, and in fact most are rarely heard to do so.

Further generalizations about bilingualism in San Mateo result from examining the various factors which tend to reinforce the stable equilibrium of bilingualism; these can be discussed in terms of factors internal and external to the community.

There are internal factors which bind the individual to the village and make it less likely that he leave. Such bondship acts as a resistance factor to bilingualism, since the social functions of the secondary language within San Mateo are quite limited.

(1) Endogamy. San Mateo is markedly endogamous. There is a fantastic web of real and fictive kinship ties which limit the individual's mobility outside the community. Failure to conform is regarded as moral aberration or psychosis, and ridicule, fines, and ostracism may result from their repeated infraction. Moreover, Huaves find difficulty in marrying into non-Huave-speaking groups; those in the immediate vicinity regard the Huaves with considerable

negative social prejudice, and the rate of Huave exogamy remains very low.

(2) The dependent nuclear family. Residence patterns reinforce the kinship bonds discussed above and the chances are likely that the nuclear family which one establishes upon marriage will be a constituent of a larger extended family.

(3) Religious and political duties. Until very recently in San Mateo, no individual physically or mentally able would have demurred from the obligations imposed by the escalafón, a series of successive public offices through which every village member must progress. Functionally, the escalafón assured the allocation of duties and leadership necessary to sustain the community; furthermore, many of the roles demand economic behavior which assures a regulated flow of goods throughout the village.

The only effective way to break these social bonds is to openly defect from the community, but the rate of defection remains low, since incentive to do so (although it is growing) is curbed by the external pressures which remain to be discussed.

There are negative external factors which tend to reinforce the features discussed above. Chief among these is the Huave's economic and political dependency on the Spanish-speaking outside world.

The Huave economy emphasizes fishing, and marine produce surpluses are destined for nearby markets. But unfortunately for the Huave, little economic power can be enjoyed since they in turn are at the complete mercy of the outside market economy for all manufactured goods and many foodstuffs, including even the staple starch in the diet, corn. And there exists on the part of the Spanish-speaking merchants a very conscious effort to keep the Huaves in this disadvantageous position.

Politically, too, San Mateo is satellite to the larger Spanish-speaking towns, principally to Tehuantepec. There is covert energy devoted to forestalling any change which will upset the present power ratio. This is manifest in the overlooking or withholding of social security services and of educational and judicial facilities, with which the township of Tehuantepec is entrusted.

The Huave is thus almost a second rate citizen. The pressures which are exerted on him make it less likely that the Huave will leave his village, and less likely that he will have to or even be able to learn Spanish. The present difficulties which the Huave encounters in attempting the transition from Indian to Mestizo are all but insurmountable.

Positive external factors are varied and contradictory. There is no consistent policy of government-sponsored education in San Mateo. Nor has the village received direct attention from any of the indigenist bureaus. The community is almost wholly illiterate, and those few individuals who are literate, are so chiefly in Spanish.

The effect of these factors is that bilingualism in San Mateo is in a state of near stable equilibrium. Physical defections from the community occur, but rarely if ever involve whole families. Defections are principally among single adult men who are socially or psychologically marginal to the community. Since these individuals rarely return, their contribution to bilingualism is negligible.

## BIBLIOGRAPHY

Caso, A. 1948. "Definición del indio y de lo indio", América Indigena 8.275-280.

Censo General 1953. Séptimo Censo General de Población 1950. Secretaría de Economía, Dirección General de Estadística. México D. F.

Diebold, A. R., Jr. 1958. Impresiones particulares sobre el bilingüismo en México. ms. New Haven.

-----1961a. Bilingualism and Biculturalism in a Huave Community. ms. Yale University, New Haven.

-----1961b. "Incipient bilingualism", Language 37.97-112.

-----1962. "A laboratory for language contact", Anthropological Linguistics 4, No. 1.41-51.

Gamio, M. 1948. *Consideraciones sobre el problema indígena*. (=Ediciones Especiales del Instituto Indigenista Interamericano, Núm. 5) México D. F.

Gillin, J. 1949. Mestizo America. In: R. Linton, ed. *Most of the World*. Columbia University Press, New York. Pp. 156-211.

Goubaud Carrera, A. 1946. "El grupo étnico indígena: criterios para su definción", *Boletín del Instituto Indigenista Nacional* 1, núm. 2-3.13-30.

Greenberg, J. H. 1956. "The measurement of linguistic diversity", *Language* 32.109-115.

Haugen, E. 1956. *Bilingualism in the Americas: A Bibliography and Research Guide*. (=Publications of the American Dialect Society, No. 26) University, Alabama.

Luís Arriola, J., ed. 1950. *Integración social en Guatemala*. (=Publicacion del Seminario de Integración Social Guatemalteca) Guatemala.

*Mapas lingüísticos de la República Mexicana*. 1944. Departamento de Asuntos Indigenas: México D. F.

Nash, M. 1957. "The multiple society in economic development: Mexico and Guatemala", *American Anthropologist*, n.s., 59.825-833.

Redfield, R., and S. Tax. 1952. General Characteristics of Present-day Mesoamerican Indian Society. In: Tax et al., op. cit., pp. 31-42.

Tax, S., et al. 1952. *Heritage of Conquest: The Ethnology of Middle America*. The Free Press: Glencoe, Illinois.

# CREOLE LANGUAGES IN THE CARIBBEAN

by William A. Stewart

It has been said of the Caribbean that the historical function of its marine expanse has been more that of a roadway to man than a barrier.[1] Perhaps it is largely for this reason that trade and commerce, closely connected with the development of a plantation economy based on sugar, spread over the area and made its island and costal land masses among the most rapidly settled and densely populated in the New World. A number of European nations, principally Spain, France, Great Britain and Holland, vied with each other for its control during the three centuries which followed Columbus' memorable landfall. As a result both of the plantation economy and of strategic settlement programs carried out by the competing European nations, the aboriginal populations of the Caribbean were soon overrun by the flow of immigrants from Europe and Africa in most places.

In many centers of settlement--especially in those established by nations other than Spain--the importation of African slaves took place on so large a scale that European cultures never really became dominant for the total population. Instead, there grew up a number of synthesis cultures which combined African, European, and occasionally Amerindian elements, to which they added innovations of their own. Where this happened, the "pure" European cultures remained restricted to the small minorities which governed and administered such settlements. Thus a cultural cleavage came about which matched the social cleavage between master and slave and, at a later date, that between Continental Europeans and the native born.

In their early stages of development, the European element in these synthesis cultures usually corresponded to those European cultures with which they were associated. For example, in a French controlled or settled area, an Afro-French culture would be likely to develop. As time went on, however, this relationship became less neat, since some areas changed hands administratively and others experienced cultural drift.

In the Curaçao island group, for example, an Afro-Spanish culture developed (possibly upon the basis of an earlier Afro-Portuguese substratum) under Spanish occupation during the sixteenth century. When the area was taken over by the Dutch in 1634, this Afro-Spanish popular culture became overlaid by the Dutch culture of the administrative élite. In Dominica, on the other hand, the predominately Amerindian population gradually gave up their Carib culture after the annexation of the island by Great Britain--but not, interestingly enough, for an English type culture, but rather for an Afro-French one, presumably because of the influence of the neighboring French islands.

This highly varied political and cultural history has given the Caribbean by far the most complex socio-political landscape of any part of the New World and causes it to stand in contrast to the rest of the Americas, where three European nations came to dominate vast areas of land, and where acculturation of the aboriginal and immigrant populations to something closely approximating the European home cultures has been, if not rapid, at least fairly thorough.

The cultural diversity of the Caribbean is perhaps nowhere more clearly manifest than in its present day linguistic make-up. At first glance, the most obvious level of linguistic diversity is that of variations in the standardized, official or "national" languages, which correspond to differences in the political affiliations of the various administrative units. The distribution of these official languages is as follows:

| LANGUAGE | DISTRIBUTION |
|---|---|
| Spanish | Cuba; Dominican Republic; Puerto Rico (with English); mainland areas of Central and South America except for British Honduras and the three Guianas; island dependencies of Venezuela.[2] |
| English | Gulf Coast of the United States and all U.S. island dependencies; Puerto Rico (with Spanish); British dependencies and members of the West Indies Federation. |
| French | Haiti; French dependencies. |
| Dutch | Dependencies of the Netherlands. |

These are the standardized languages which are used for all official and representative purposes, and which are usually found given as the languages "spoken" by the particular populations. For those areas where Spanish is the official language, it is indeed the language used by virtually everyone for practically every purpose. In the other cases, however, it takes very little first hand observation for the visitor to discover that in many areas there exists a clear linguistic dichotomy in that beside the official language there is used another, obviously different form of speech. This other form of speech is in most cases the really popular medium of communication for the area, and is used for practically all unofficial and non-representative purposes.

There are several such popular languages in the Caribbean--as many, in fact, as there are official languages. Further, these popular languages show a definite historical relationship with the four predominant official standard languages of the area in that their vocabularies are obviously derived from those of the standard languages (or from Portuguese, which is no longer used in the Caribbean region). Yet the geographical distribution of these popular languages is by no means identical with that of their standardized relatives, and they are different enough from them in other respects--particularly in grammatical structure--to leave little doubt that they qualify as independent languages. Viewed apart, these popular languages are "new languages" in that they show evidence of having acquired their present shapes much later than their nearest related standard languages. These new languages, which are clearly related to European languages, and yet have grown up entirely outside of Europe, have been given the name of Creoles.[3]

The most widely used Caribbean Creole by far is that with a vocabulary based upon French.[4] Used over an area extending from one end of the Caribbean to the other, the four major dialects of French-based Creole, i.e. those of Louisiana, Haiti, the Lesser Antilles, and French Guiana, are all mutually intelligible. French-based Creoles are generally known by the name créole or patois in areas where they are spoken. The second most widely used variety is a Spanish/Portuguese-based Creole known as Papiamentu. Thirdly, there are two main varieties of English-based Creole, which are known as Sranan and Saramakkan.[5] Lastly, there is one Dutch-based Creole, usually called Negerhollands (i.e. Negro Dutch), which has now become practically extinct, but which once had enough native speakers to justify the translation into it of large portions of the Bible by Danish missionaries.

The distribution of these Creoles is shown by the following table:

| CREOLE & ESTIMATED NUMBER OF SPEAKERS | AREA | POLITICAL AFFILIATION |
|---|---|---|
| French-based Creole (4,500,000) | Louisiana | (U.S.) |
| | Haiti | (Indep.) |
| | Lesser Antilles | |
| | Guadeloupe | (Fr.) |
| | Les Saintes | (Fr.) |
| | Marie-Galante | (Fr.) |
| | Martinique | (Fr.) |
| | Dominica | (G.B.) |
| | Trinidad | (G.B.) |
| | Grenada | (G.B.) |
| | St. Lucia | (G.B.) |
| | French Guiana | (Fr.) |
| Spanish/Portuguese-based Creole (200,000) | Curaçao | (Neth.) |
| | Aruba | (Neth.) |
| | Bonaire | (Neth.) |
| English-based Creole (80,000) | Surinam | (Neth.) |
| Dutch-based Creole (negligible) | Virgin Islands | (U.S.) |

A comparison of the political affiliation of each of the areas where a particular Creole is spoken with the distribution of Caribbean official languages indicated earlier will show that in many cases the Creole and the official language spoken in a particular locality may not be related. Indeed, largely due to changes in the political affiliation of many areas after the Creoles had become established, French-based Creoles, besides being spoken in French dependencies and in Haiti (formerly the French colony of St.-Domingue), are spoken in areas where the official language is English, i.e. in the State of Louisiana and in a number of British dependencies. Also, English-based and Spanish/Portuguese-based Creoles are spoken in dependencies of the Netherlands, where the official language is Dutch. The Dutch-based Creole, however, has been centered in the Virgin Islands, where the administrative language was once Danish and is now English.

The population in a Creole speaking area can be divided into three different categories of speakers on the basis of whether individuals know the Creole, or the Standard, or both:[6]

1. Monolingual Creole speakers (i.e.those who have a working knowledge only of the Creole)
2. Monolingual Standard speakers (i.e. those who have a working knowledge only of the Standard)
3. Bilinguals (i.e. those who have a working knowledge of both the Creole and the Standard)

The first group consists for the most part of native inhabitants of the area who have had little or no formal education of any kind. The second group consists almost invariably of persons born and raised outside the area (e.g. immigrants, colonial administrators, high churchmen, etc.) who know the Standard either as their mother tongue or as a second language, but who have never become conversant with the Creole. The third group consists in the main of those native inhabitants who, in addition to their Creole, have learned the Standard at home or in school.

The relative numerical weight of the three groups varies from area to area, but the number of monolingual Standard speakers is always fairly low. As would be expected, they and bilinguals are more common in urban centers than in rural localities.

One of the most interesting aspects of the interrelationship between Creoles and Standards in the Caribbean is their relative functional distribution in the national communicative network, and the different sets of attitudes which the society attaches to them.[7] One technical fact which affects the functional difference between Creoles and Standards in the Caribbean is that Creoles are seldom written, so that the Standard is the usual language to employ for all normal writing purposes in national life, e.g. signs, correspondence, newspapers, etc. However, the functional difference between Creoles and Standards goes beyond writing. Even in purely spoken situations, each language has its appropriate uses, which are largely correlated with the particular social situation. Of all the contextual variables in the social situation, there are two which appear to be especially important as determinants of whether the one or the other language is the more appropriate one to use. These are:

1. Public (i.e. impersonal or representative) vs. private (i.e. personal and non-representative) behavior.

2. Formal (i.e. prescribed) vs. informal (i.e. non-prescribed) behavior.

The functional distribution of the two languages in terms of these variables can be diagrammed as follows:

| | FORMAL | INFORMAL |
|---|---|---|
| PUBLIC | Standard | Creole<br>(Standard) |
| PRIVATE | Standard<br>(Creole) | Creole |

Examples of public-formal activity are official governmental activities, legal procedures, academic and other formal educational activities, public speaking, the programmed part of radio and television broadcasts, and ceremonies of introduction between strangers. Examples of public-informal activity are the singing of popular songs, the telling of anecdotes, bartering, advertising, and general conversation between acquaintances when in the presence of strangers. Examples of private-formal activity include unofficial government connected activities, formal receptions and parties, and general conversations between strangers. Finally, examples of private-informal activity will include practically all cases of interaction which are exclusively between friends, lovers, or relatives, such as within the family, at informal parties, in the street, at work and at play, etc.

In those cases where the chart indicates that either language may be employed, usage varies locally with the tendency in the direction of the language listed first. In Haiti, Curaçao, and Surinam, for example, the tendency is to use the Creole, while in Louisiana and the French and British dependencies, it is to use the Standard.[8] It will be noted, however, that for certain kinds of activity only one of the two languages is appropriate, and consequently that neither language is used for all situations. Since monolinguals tend to be excluded from direct participation in those activities where the language used is unknown to them, it is clear that only those who are bilingual will enjoy full integration into the national communicative network. For those who are monolingual speakers of the Standard, this situation does not present very serious problems, since use of the Standard by foreigners in private and informal situations tends to be tolerated and even expected. Such people are inconvenienced primarily by their inability to communicate with monolingual Creole speakers.

Far more serious are the consequences which arise for the monolingual Creole speakers, for the language distribution excludes them from direct participation in public-formal activities, and this includes important administrative and legal procedures. In areas where the percentage of monolingual Creole speakers is high, this matter of differences in language function may give rise to grave problems of political and legal non-representation.

The functional distribution of the Creole and the Standard as outlined above holds for practically all of the Creole speaking Caribbean. However, in cases where the Creole and the Standard have a close lexical relationship, the two languages may be much more closely identified with each other, with the result that the functional distribution of the two languages may be even more refined. In such cases, the two languages may even come to function, in the discourse of bilinguals, much as if they were but different style levels of the same language. This special kind of language relationship--

sometimes called diglossia, to distinguish it from other kinds of bilingualism--exists between the Creole and the Standard in Haiti and the French dependencies.[9]

Interwoven with these largely unconscious and automatic conventions governing the functional distribution of Creoles and Standards, there exists almost everywhere a set of conscious social attitudes toward them, by means of which the society expresses differences in the degree of prestige which it attaches to each of the two kinds of language. One of the basic reasons why the languages are ranked differently is that Creole speakers--and this includes monolinguals--can be just as aware as anyone else that there is something different about Creoles as compared to other languages. There are several reasons why such an awareness may exist:

1. It is observed that Creoles are seldom written, and are never used for formal, public, representative, or any other "important" purpose by persons who have responsibility and authority in the society.

2. It is observed that Europeans (who have generally high prestige in the Caribbean) do not speak any of the Creoles in their home countries.

3. Where speakers of a Creole come into contact with the lexically related Standard, it is almost unavoidable that there will be a comparison of the two related forms of speech. When this happens, the political and cultural prestige and linguistic formalism attached to the Standard will almost invariably weigh in the Standard's favor, and differences from it observed in the Creole will be regarded as substandard.[10]

4. The history of the Caribbean is recent enough so that there is a generally shared knowledge that development of the Creoles has some connection with the slave trade and early plantation economy.

As a result of this awareness, there is a general tendency to "play down" the role of Creole in the life of the area. In official publications, for example, the number of speakers of the Standard is almost always overestimated and, where the Creole and the Standard are related, the temptation to classify them both under the language designation of the Standard alone often proves irresistable. In addition, bilinguals tend to be curiously inaccurate in their vocal estimates of the extent to which they use one or the other language, and the error is always in favor of the Standard. Thus individuals can be heard to state something like "Oh, I never speak Creole; I am from an educated family and the only language which I am used to speaking is Standard", when in fact direct observation will show that very person using both the Creole and the Standard, along the lines of the functional distribution indicated earlier.

In addition to such generalized attitudes, there exists in most areas a set of clichés which deprecate the Creole and thereby imply the superiority of the Standard. These clichés are so frequently parroted that it is apparent that they are not simply the product of amateurish linguistic speculation on the part of individuals, but rather are formal tenets of the culture. Like any such tenets, they are simply assumptions which are acquired and passed on unquestioningly--in the home, in the street, and in this case most of all, in the school--with no attempt made by most people to judge their validity by empirical observation. Indeed, the questioning of any of these clichés by a member of the local society will be very likely to be regarded by the other members as a somewhat antisocial act.

It should not be assumed, however, that these linguistic clichés have an entirely mythical basis. Given certain facts, they are in many cases just the kinds of conclusions which one would be expected to arrive at in the absence of scientific techniques of linguistic analysis. In other ways, they may well represent the normal social differentiation which can be expected between a standardized and a non-standardized language when both exist together.[11]

Since these clichés and the attitudes which they express undoubtedly have their effect on local administrative policy with respect to language problems, and since every serious student of the area can expect to encounter them at some time, it is well worth while to review them here:

Creole is not a language. The rationale of this cliché is usually the fact that Creoles are not standardized, not generally written, and have little in the way of an important serious literature. Although it is quite possible that standardization, writing, and literature may have important social implications for a language, they are not in themselves scientifically acceptable as criteria for defining what is and what is not a real language.

*Creole has no grammar; it is a "made up" form of speech*. All languages have grammatical structure, Creoles no less than other languages; moreover, much of the grammar of most Caribbean Creoles has already been described. However, since knowledge of these descriptions is still limited largely to linguistic specialists and has not become accepted as formalized conventions in the cultures of the Creole speaking areas, it is understandable that a speaker of Creole (whose grammar has been internalized unconsciously in childhood) may get the impression that he is "making up" the language as he goes along.

*Creole is a corruption of [the lexically related] Standard*. When two closely related languages occur side by side, one with a grammatical structure which has been formalized into known and accepted conventions, and the other with one which has not, there is a tendency to regard departures in the behavior of the latter from the conventions of the former as a departure from structure itself. Thus, a Creole may sound to those who know the lexically related Standard as a corrupted or badly spoken version of it. What has really happened to the Creole, however, is that it has simply developed its own behavioral patterns, and is no more a "corruption" of a Standard with which it happens to have a historical relationship than, say, French is a corruption of Latin merely because it has developed out of Latin into a different form of speech.

*Creole has no uniformity of usage*. Persons who are predominantly speakers of a Creole can be expected to be more sensitive to variations in that Creole than in a Standard which they know only from formal samples. There is, nevertheless, an unjustifiable degree of exaggeration of the amount of regional variation for Creoles as compared with Standards. In Haiti, for example, people are fond of pointing out one of the few clear cases of regional lexical variation in Creole, i.e. *kanistè(r)* 'tin can' in the North for *ma(r)mit* or *fè(r)blã* elsewhere. Yet they seldom call attention to like differences between their French and that of the Continent, e.g. *gasoline* in Haiti where France uses *essence*. A certain amount of regional variation is natural in any widely used language, although one would expect it to be less marked in a standardized language than in one which remains unstandardized.

*Creole is a mixture of many languages*. Some Creoles, e.g. Papiamentu, may indeed show evidences of a vocabulary that is more mixed than that of their closest lexically related Standards. But other Creoles, such as the French-based ones, have vocabularies which are overwhelmingly based on one and the same source language. It is interesting to note that the French-based Creoles have less vocabulary of non-French origin than Standard English has of vocabulary of non-Germanic origin. The fact is that all languages engage in at least some borrowing of vocabulary from other languages, and Standards do this no less than Creoles.

Similar to the foregoing cliches both in tone and implication is the traditional account of how the Creoles came into existence. This account holds that the Creoles originated in colonial times as linguistic compromises resulting from an incomplete acquisition of the Standards by the newly imported African slaves, whose linguistic mistakes were reinforced by deliberate efforts on the part of their European masters to "talk down" to them. Since the contact between master and slave occured for the most part on plantations, it is assumed that the Creoles were purely local developments.

Since this theory is not entirely unreasonable, and in fact ties in with Caribbean history rather neatly, it is easy to see how it has gained wide acceptance. However, more recent linguistic studies of the Caribbean Creoles have highlighted new kinds of evidence which suggest that the scope of their historical development has been greatly underestimated.

In the first place, none of the Caribbean Creoles are now mutually intelligible with their lexically related Standards, and if early accounts are to be trusted they apparently never were. Since this is the case in spite of the considerable lexical overlap between a Creole and its related Standard, the principal reason for the lack of intelligibility would seem to lie in the fact that their respective grammatical structures differ greatly. This discrepancy--similar vocabularies on the one hand and dissimilar grammars on the other--has never been satisfactorily accounted for by the traditional version of the development of Creoles.

Within a given Creole, however, dialectal variation is usually so moderate that mutual intelligibility over different dialects is the rule rather than the exception, even for cases where dialects may be geographically separated from each other. The most striking example of this occurs in the French-based Creole, where speakers from opposite ends of the Caribbean --from Louisiana in the north and French Guiana in the south-- are able to converse together with a minimum of misunderstanding. Similar cases, though on a much smaller scale, are common in the other Creoles as well.[12] Although this mutual intelligibility is in part due to the fact that the various dialects of a Creole derive their vocabularies from the same source language, the fact that Creoles are not mutually intelligible with their lexically related Standards shows that vocabulary alone is not a sufficient explanation. What is of equal importance for the Creoles is the fact that their like vocabularies are matched by grammatical structures which are also very much alike. Since it is highly improbable that dialects showing such overall similarities would have developed quite independently of each other, it is difficult to avoid the conclusion that the dialects of at least the French-based Creole must have had some common origin before they became implanted in widely separated parts of the Caribbean. This conclusion receives some support from French-Creole texts dating from not much more than a half-century after the first French settlements were established, showing that the language had by that early date already evolved into something fairly close to its present form.[13]

A very likely explanation of this situation is to be found in a kind of linguistic development which involves rapid grammatical shift within a language, and which has been attested to in the case of a number of specialized trade languages which have developed in various parts of the world. In this type of change a language undergoes a restructuring of its basic grammar while preserving its basic vocabulary relatively intact. Languages which have undergone such a restructuring process are usually called pidgin languages, and the process itself is often referred to as pidginization.[14] Accordingly, it is quite possible that the Caribbean Creoles first developed in their present shape as pidginized forms of the respective European languages--presumably at some date before colonization was over, if the high level of dialectal uniformity within each Creole is to be explained--and gradually became creoles as they acquired native speakers.

Such a hypothesis certainly helps to explain both the grammatical gap which exists between the Creoles and their nearest European relatives, as well as the high degree of linguistic uniformity within each individual Creole. However, it has also been noted that the Caribbean Creoles exhibit an impressive degree of correspondence in their respective basic grammars. This correspondence includes not only similarities in syntax but even cases where specific grammatical morphemes are shared in common.[15] Furthermore, some investigators have shown that there is a similar type of correspondence between the Caribbean Creoles in general and certain pidgin and creole languages of Asia and Africa.[16] This strongly suggests that there is some kind of genetic relationship between the various Caribbean Creoles, and even between them and other creole-type languages spoken outside the Caribbean. The nature of this relationship is suggested by evidence that at least some of the Caribbean Creoles have undergone a kind of lexical shift, whereby the vocabulary derived from one source language has been largely replaced, through a process of widespread borrowing, by a more recent vocabulary derived from another language, while the original grammatical structure is preserved practically unchanged.[17] This process of relexification seems to be the converse of restructuralization. For example, if a language A can be shown to derive its vocabulary from language B and its grammatical structure from language C, then language A can be both "restructured B" and "relexified C" at the same time.[18] Thus, although the Caribbean Creoles are usually considered (from the point of view of their vocabularies) to be restructured or pidginized forms of French, etc., it is just as possible to consider them all (from the point of view of their similar basic grammars) as relexified forms of some prior language. But then the question is, what prior language? It would have to be a language which had a historical connection with the settlement of the Caribbean, one with a basic grammatical structure like that of the Caribbean Creoles, and one toward which other kinds of linguistic evidence, such as cases of a substratum vocabulary, would also point.

The clearest cases of lexical substrata in the Caribbean Creoles appear to be of Portuguese origin. Therefore a very likely source language for the Caribbean Creoles would be the Portuguese trade pidgin which apparently originated as early as the middle of the fifteenth century and was once spread throughout areas of Portuguese contact and expansion in Africa and Asia.[19] This was the language which was apparently used between the Portuguese and West Africans, and which became the earliest common language used in the slave trade.

When the French, Spanish, English, and Dutch entered into the trade, it is probable that this pidgin served as their first language of wider communication, both with the Portuguese and with the West Africans. Subsequent relexification of the pidgin could have taken place both at the slave factories on the African Coast and, in some cases, later within the Caribbean itself. The latter cases are in all likelihood the ones which have furnished documented evidence of relexification. These new French, Spanish, English, and Dutch pidgins were then used as the primary means of master-slave communication in the new plantation life in the Caribbean settlements, and were the immediate ancestors of the modern Creoles.

It is known that there was a deliberate policy in almost all of the early Caribbean settlements of mixing the slaves linguistically. This encouraged learning of the pidgins, which then became the only means of communication both between Europeans and Africans as well as between Africans with different linguistic backgrounds. By the time immigration to the Caribbean from Europe and Africa had begun to ebb, these contact languages had become well established as the mother tongues of those ensuing generations born and brought up in the new land.

For well over two centuries, Creoles and Standards have been used side by side in the Caribbean. Reflecting a traditionally rigid social structure, the relationship between Creoles and Standards has been one of a well defined, mutually exclusive social and functional distribution. As long as Caribbean society remained essentially an extension of colonial plantation social structure, little motivation existed for equalizing the linguistic opportunities for participation in local cultural and administrative affairs. More recently, however, social and technological change has begun to penetrate the region, and this has prompted among other things the adoption of policies aimed at reducing those linguistic differences which it is felt promote social and political inequality, i.e. the contrast between literate and illiterate, and between use of the Creole and of the Standard.

Policies dealing with multilingualism are usually directed toward bringing about the eventual use of a particular language by all members of the society. Ideally, this implies complete monolingualism in the language chosen for use. Considering the standardization and high prestige of the official languages in the Caribbean, it is not surprising that their elimination is never seriously considered, even though the Standard may be known by less than ten per cent of the population, as is the case in Haiti. This leaves the discouragement of the Creole or else generalized bilingualism as the only practical alternatives.

In some areas of the Caribbean the Creole, once widely used, has been largely replaced by the Standard within the past one hundred years. For example, in Louisiana, the Virgin Islands, Grenada, and Trinidad, use of the Creole is becoming more and more restricted to members of the older generation. Where a well marked trend in this direction already exists, the achievement of monolingualism in the Standard is a definite possibility, and could probably be speeded up by an enlightened program encouraging exclusive use of the Standard and making easier the transition to it from the Creole.

However, in areas where the Creole is the primary medium of communication, as in Haiti and Surinam for example, it is doubtful if a similar policy would ever meet with much success. Indeed, it would probably only frustrate communication further by discouraging the use of Creole in situations where the Standard is not a realistic substitute. In view of the general reluctance to do away with the Standard, the only realistic course available for dealing with such a situation would seem to be one which would lead toward general bilingualism.

In any event, the selection of the most suitable linguistic policy for a given area in which both a Creole and a Standard are used should be made only after taking into account the relative importance of the two languages. This should be determined wherever possible by careful study carried out by competent linguists or social scientists.

In situations where the encouragement of bilingualism appears to be either preferable or necessary--especially in cases where some speakers of the Creole are not likely to become involved in situations which would require the Standard (e.g. the Saramakkan speakers in Surinam)--it might be useful to impart to the Creole through the process of standardization some of the status and versatility enjoyed by the Standard.

In the past there have been attempts to standardize certain Creoles to some extent, but these have never really caught on. One of the main reasons of failure seems to be that normalization has in most cases not been carried beyond the devising of an orthography. The reduction of a language to writing without a corresponding formalization of other aspects of the language is simply a meaningless step in terms of what most people expect in a prestige language.

In future programs of standardization for Creoles, an effort should be made to present the users of the language with a set of conventions dealing not only with orthography but with grammatical and lexical usage as well. Before such standardization is achieved, however, there should be no attempt to spread the use of the Creole into the functional distributional area of the Standard, since the use of a language outside of its prescribed function without an accompanying change in its status is likely to be considered locally as inappropriate or even ludicrous, and consequently might meet with some resistance.

Another matter related to the use of a Creole and a Standard together concerns their respective roles in local programs of second-language teaching. It would seem obvious that the structure of the Creole should be taken into account in teaching the Standard, yet textbooks for teaching Caribbean Standards are usually copies of European works designed primarily for children who are monolinguals in the language. Some very important work needs to be done in the development of pedagogical materials which contrast the differences in linguistic structures of a Creole and Standard.

Finally, in teaching second languages to those who are already bilingual in Creole and the Standard, there is the question whether in teaching the foreign language the linguistic materials should be designed on the assumption of major interference from the Standard, from the Creole, or from both. But scarcely any attention has been paid to this important matter. In teaching English and Spanish in Haiti, for example, both local and foreign pedagogical materials utilized have all been of the same kind as those used for teaching English and Spanish to monolingual speakers of French. Yet anyone who hears Haitians speak English or Spanish will quickly notice that they speak these languages with an accent which is noticeably different from that of a Frenchman, and that they also appear to have special kinds of difficulties with grammar as well. Much more attention should be devoted to orienting foreign language teaching in Creole speaking areas to an approach in which greater account is taken of potential interference from the Creole.

NOTES

1. The Caribbean area proper is usually taken as that covered by the Greater and Lesser Antilles. By extension, however, the term is often made to include certain other regions which share many of the topological, historical, and cultural characteristics of the Antilles, i.e. parts of the North American Gulf Coast, the Bahamas, and the Central and South American costal regions down as far as and including the three Guianas. Where it is necessary to make a distinction, the marginal area is referred to technically as the Circum-Caribbean. In this paper, the term Caribbean will always be used in its most inclusive sense.

2. The political affiliation of the various island and mainland forms will be found indicated (by color coding in most cases) on any good map of the area.

3. The term Creole, as applied to a language, should not be confused with other uses of the term. In particular, it should be emphasized that a Creole language is not necessarily simply a language spoken by people who are known in one or another locality as Creoles--whatever these may be. Here the term is used to distinguish languages which have certain characteristics from other languages which do not have those characteristics. Anyone may be the native speaker of a Creole language, providing he has been brought up in an area where one is spoken.

4. Creoles will be identified in terms of the origin of their vocabulary. Thus a Creole language with a vocabulary which is predominantly of French origin is termed a French-based Creole, and so on. This is the common practice, although it will be seen later that there are other possible ways of assigning genetic relationship for Creole languages.

5. There are in addition local varieties of English which are spoken in Jamaica, British Honduras, Trinidad, British

Guiana, and the Lesser Antilles. These, however, are not regarded as Creole languages in this survey. The reason is that when all the varieties of English of any one of these areas are taken together, the result is a continuum of variation, with Standard English at one end and the most deviant local form at the other. This is reminiscent of usual dialect variation around a standardized norm, and contrasts with the situation for the Creoles where, no matter how much they may be compromised in the direction of their lexically related European standard languages, there is always a structural gap between the two at some point, and consequently it is always clear just which language is being spoken at any particular time. What this means is that there is some constant criterion available for distinguishing Creoles from related standard languages, whereas Jamaican and other regional varieties of English are best treated as dialects of English. That such dialects are often referred to locally as *creolese* (particularly in the Lesser Antilles) should not be allowed to obscure their basic difference from real Creoles. Also, it is worth noting that the Gullah dialect, spoken on the islands off the coast of South Carolina and Georgia, is possibly another English-based Creole. It has not been included in this survey, however, since its geographical distribution puts it clearly outside of the Caribbean, and because its relation to Standard English and to the other English-based Creoles needs further study and clarification.

6. Since much of this survey will consist of statements about the general characteristics of Creoles as a class of languages and their relation to related European standardized languages, the term *Creole* will refer hereafter to any or all Creole languages, and *Standard* will refer to any or all of the standardized European languages which have an official function in the Caribbean.

7. The relationship of Creole and French in Haiti is in all probability the case for which the most sociolinguistic descriptive material is available and, with very few modifications, it seems to be typical of most of the Caribbean. For an overview of the situation in Haiti see the chapter on Creole in James G. Leyburn, *The Haitian People*, New Haven, Yale University Press, 1941, and the insightful article by Edith Effron, "French and Creole Patois in Haiti", *Caribbean Quarterly* 3:14.199-213 (1954). For information on one other French-Creole area, see Mervin C. Alleynde, "Language and Society in St. Lucia", *Caribbean Studies* 1:1.1-10 (1961).

8. In Louisiana, French exists beside English as a Standard, though in an unofficial capacity and with steadily diminishing use. The point is that Louisiana Creole speakers will have a tendency to employ some sort of Standard in the situations indicated.

9. For a general description of diglossia, see Charles A. Feguson, "Diglossia", Word 15.325-340 (1959). The mechanism of diglossia in Haiti has been described in William A. Stewart, "The Functional Distribution of Creole and French in Haiti" (to appear in Georgetown University Monograph Series on Languages and Linguistics, No. 15, Thirteenth Annual Round Table of Languages and Linguistics).

10. The opportunity to compare a Creole with a lexically related Standard is probably an extremely important factor in determining the kinds of social attitudes which will develop toward a Creole. It is interesting to note that in Surinam, where the speakers of the English-based Creoles have had only slight contact with Standard English, the Creoles have acquired some degree of social prestige.

11. It is striking that these clichés about Caribbean Creoles are almost word for word counterparts of what was once said in Europe about the Standards themselves, back when these were as yet unstandardized and occupied an inferior social position with respect to standardized and written Latin.

12. By exception, the two main varieties of English-based Creole, Sranan and Saramakkan, are not mutually intelligible, though there is general intelligibility within each of these.

13. Examples of most of the important early French-Creole texts are given in Elodie Jourdain, Du francais aux parlers créoles, Paris, 1956.

14. See Robert A. Hall, Jr., "Pidgin Languages", Scientific American, February 1959, pp. 124-134. In most cases, restructuring has been attributable to the influence of the grammatical systems of other languages, or to certain functional pressures for maximum grammactical simplicity.

15. For technical details, see the two articles by Douglas Taylor, "Language Shift or Changing Relationship?",

International Journal of American Linguistics 26.155-161, (1960), and "New Languages for Old in the West Indies", Comparative Studies in Society and History 3.277-288 (1961). Although the term basic grammar refers essentially to such fundamental structural characteristics of the language as verbal categories, syntactic structures, etc., it is interesting to note that in Creoles certain kinds of lexical items tend to be associated more with the grammar than with the vocabulary, insofar as historical derivation is concerned. These include grammatical morphemes such as prepositions, verbal particles, and some pronouns. Also behaving this way are certain types of symbolic behavior which are marginal to language, e.g. vocal qualifiers and ideophones.

16. See R. W. Thompson, "A Note on Some Possible Affinities Between the Creole Dialects of the Old World and Those of the New" in R. B. Le Page (ed.), Creole Language Studies, II, London, 1961, pp. 107-113. A pidgin language is a simplified version of some other language, and usually develops in certain contact situations when no language is previously known in common. Pidgins are definitionally limited in scope to languages which have no native speakers. Once native speakers develop for a pidgin, it becomes a creole language.

17. There is some documentation of this process in Papiamentu and in Sranan; see the two articles by Douglas Taylor noted in footnote 15, above.

18. Languages which exhibit a dichotomous genetic relationship of this type might be termed heterogenetic languages.

19. Languages which seem to be direct descendants of that Portuguese pidgin are presently spoken as creoles in certain islands off the coast of West Africa, e.g. Cape Verde, Annobón, and São Tomé.

# LINGUA FRANCAS, WITH SPECIAL REFERENCE TO AFRICA

by William J. Samarin

## Origin of the term

The language of the Crusaders was not the first lingua franca in the history of mankind, but it furnished the name for all such similar languages ever since.

The horde which descended upon the Muslims on the eastern shores of the Mediterranean had come from many parts of western Europe. The native languages (the vernaculars) of these priests and soldiers, merchants and blacksmiths, porters and page boys, were forerunners of English, French, Italian, and German. Though Latin was the language of religion and of learning, the Crusaders did not all speak Latin, nor did they speak each other's languages. While there were certainly bilinguals among them, there were with equal certainty people who did not understand even other dialects of their own language: an Englishman from the South might find it difficult if not impossible to understand an Englishman from the North on the first encounter.

Linguistic diversity (i.e. multilingualism) has always set the stage for the development of lingua francas, and the multilingual Crusaders found a need for such a language. They found it in the language of Provencal, spoken along the southern shores of Europe between Marseilles and Genoa, and this language became the basis of a language used by the Crusaders all the length of the Mediterranean. The 'French language' (which is what lingua franca literally means, though it was most likely only a particular dialect of a Romance language) thus came to be used among the Crusaders and with the non-French speaking peoples who had learned this language.

The use of this particular form of French eventually died out, but not before having left its own name as its heritage for languages which are used in a similar fashion. Today when we find a language which is commonly used by people whose native languages are different, we describe it as a lingua franca. One can even speak of a lingua franca for a particular occasion. If, for example, a Norwegian and a Persian, whose languages were unknown to each other, talked in English, this language would function as their lingua franca.

## A definition

If the term lingua franca is to be used with the widest possible meaning it is necessary to adopt a definition with the fewest restrictions. The following (UNESCO 1953, p. 46) is worthy of general adoption: "A language which is used habitually by people whose mother tongues are different in order to facilitate communication between them." An additional requirement included in some definitions is that a lingua franca be a pidgin language (a so-called "hybrid" or "mixed" language), but it is not true that all lingua francas are pidgins. A second requirement sometimes found is that a lingua franca be used for commercial purposes, but this is again not always true of lingua francas, though they frequently arise in a commercial environment.

Other terms compete with lingua franca as designations for the kind of language being discussed.

1. Trade language ("langue de traite") is usually used for some language not included among the world's majority languages and which is used by some people as a second language in commercial situations. All trade languages are therefore lingua francas, e.g. Kituba and Hausa in Africa.

2. **Contact language (probably equivalent to French "langue véhiculaire") is a lingua franca whose use is not necessarily habitual. It is the most neutral of the terms.**

3. International (or Universal) language is a lingua franca whose use is actually or virtually international. Some writers, however, make the error of using it of lingua francas indiscriminately, even of those with restricted use, e.g. English, French.

4. Auxiliary language is generally meant to describe an artificially devised lingua franca, e.g. Esperanto.

## Kinds of lingua francas

Since a lingua franca is simply a language used to communicate across linguistic barriers, it can itself be any kind of a language; natural, pidginized, creole, or artificial. By natural language is meant any language acquired by the normal processes of enculturation. Natural languages are thus the mother or native languages of some people. When a natural language is acquired as the second language of different people, it becomes their lingua franca. In the process of becoming a lingua franca, a language often loses some of its vocabulary or is simplified in its phonology or grammar. Where the structure of the language suffers serious modification, it is said to be pidginized.

A pidginized language (or simply pidgin) traces its lineage to at least one natural language. A pidgin, strictly speaking, is not a natural language (see above). However, it may become a natural language; at this point it may be called a creole.

Unlike either of the preceding kinds of lingua francas are the artificial languages. Not all such languages are lingua francas, for some are intended for very restricted use. All such languages were created ad hoc, though based on some natural languages, for express use as lingua francas.

## Natural language lingua francas

For Sub-Saharan Africa the more important natural language lingua francas--outside of English and French--are the following.

1. Swahili, a Niger-Congo language spoken in Kenya, Tanganyika, and much of Uganda.

2. Lingala, a Niger-Congo language spoken in the Congo.

3. Various other Niger-Congo languages, such as Bambara and Fulani in the West; Bulu in the Cameroons; Ngbandi in the Northern Congo.

4. Hausa, a Chadic language of the Afro-Asiatic family, used as a trade language in Dahomey, Togo, Ghana, Upper Volta, Nigeria, and the Cameroons.

## Pidgin languages

Pidgin languages as a group typify most completely the lingua francas in their development and structure. Bona fide pidgins are found all over the world and have been under observation for about one hundred years. However, it was not until recently that they have been submitted to careful analysis. The reason for this was the opinion that pidgins at their best were nothing more than poor approximations to cultural languages such as English and French.

However, what people failed to see was that the pidgins were not mixtures of linguistic debris, but rather examples of genuine linguistic structures. The earliest of such studies was Hall's of Melanesian Pidgin English (Hall 1943).

The bibliography reveals that today many scholars are working on several different pidgins, and the published literature has not only made available working data on the languages, but has also raised several interesting theoretical or methodological questions, such as the following: (1) Can pidgins be classified as belonging to language families like natural languages? (2) What are the psychological and social correlates of their use? (3) Can they be described with the same rigor as natural languages?

Following is a list of African pidgins (Niger-Congo) languages.

Bangala-Lingala: spoken in the Congo and derived from Ngala; Bangala is supposed to be the more simplified of the two.

Bulu-Yaounde: two dialects of the same Cameroon language; although Bulu is a natural language lingua franca, there appears to be a pidginized form of it used by--among others--truck drivers.

Fanagalo (also known as Fanikalo, Chilapalapa, Kitchen or Mine Kaffir, Isikula, Cikabanga, Cilololo, Cilungubo: "Basic Bantu"): a pidgin Zulu of Southeast Africa.

Kituba (also known as Commercial Kikongo, from which it is derived, Simple Kikongo, Kibulamatadi, Kikwango): spoken in the Congo.

Sango: derived from a language of the same name which is an intimate dialect with Ngbandi; spoken in the Central African Republic.

Swahili: derived from the language of the same name; spoken in Tanganyika and neighboring areas; the dialect of the Eastern Congo is called Kingwana.

As linguistic curios a few others might be cited: Coromanti, spoken in Jamaica around 1800, and derived from some West African language (Le Page 1960, p. 98); pidginized Ewondo ("Ewondo populaire") in the southern Cameroons, (Pierre H. Alexandre in correspondence); Nago, probably based on Yoruba and ultimately used only in certain Brazilian pseudo-African cults (Pierson 1942, p. 73).

If a pidgin language is used long enough in any one area, children reared in the area acquire it as their native language. When this happens the pidgin becomes the natural language of a segment of the population and is said to be creolized. This creolized language may outlive the competing languages or it may become fully naturalized. At this point it is said to be a creole. (While the process here described is well known, the terminology may not represent a consensus; writers differ in the technical use of the word "creole".)

The list of pidgin lingua francas scattered throughout the world hides an important fact: though numerous and widespread, most of them are based on some Indo-European language--Dutch, English, French, Portuguese, or Spanish. When located on a map, another significant fact appears: every one is located adjacent to a marine expanse. This suggests that the history of these pidgins is somehow connected with oceanic travel, which is in fact the case. What is not obvious, however, but certainly demonstrable, is that a remarkable number of these lingua francas owe their origin to the heroic masters of the Age of Exploration, the Portuguese. In the fifteenth century there developed a pidgin Portuguese which may have originated in the first contacts with the Africans, but ultimately spread to the ports of the Far East (Lopes da Silva 1958). According to Whinnom, "Even at the beginning of the eighteenth century, English trade at Canton was carried on through Eurasian Portuguese interpreters who translated the Cantonese into a Portuguese pidgin comprehensible to the English sailors" (Whinnom 1956, p. 7, fn. 13).

The influence of Portuguese pidgin on the other Atlantic pidgins was effected through the slaves for whom the language had become a lingua franca as early as the fifteenth century (the vestiges of this Portuguese sub- or ad-stratum has been documented for several languages). This language made communication possible among the slaves and with their masters until another language was learned. When this second language was pidginized, it often was pidginized along the pattern of the earlier one. For this reason Loftman (1953) speaks of a "Caribbean creole" grammar, regardless of the Indo-European language superimposed on it.

The origin of pidgin lingua francas are for the most part hidden behind the veil of history. In only brief sketches can the origin of some be described. The past of others perhaps is buried in obscure or unknown records awaiting the scholar.

More important than knowledge of the origin of these lingua francas, however, is that of their subsequent history. What is significant about their linguistic development, about their part in the lives of human beings? What then are some of the linguistic and sociological concomitants in their use?

Linguistic changes

All lingua francas undergo certain linguistic changes which in their extreme form are called pidginization. As the use of a natural language is taken up by people for whom it is only a second language, either in the same or in a different area, it suffers from linguistic interference, which is common in all bilingual situations. The speakers of the first (or source, S) language receive the second (or target, T) through the grid of their S language. If, for example, the S language had no gender whereas the T language did, one finds--as in the case of Hausa feminine--that one is lost. In the same way other features of the language are "reinterpreted."

This process might be called simplification, and it affects the grammatical structure as well. Many of the devices used for distinguishing various parts of sentences from one another and for showing the relationships between these parts may be lost. In Ngbandi, a tone language with three levels of tone, the pronouns and verbs are intimately related in the aspectual system: different combinations of pronouns and verbs are used to mark aspect and tense. Tone is also used to derive nouns from verbs, and in other parts of the language varies considerably, but regularly, in the context of other tones. Practically all of this has been lost in Sango, the contact language derived from it. While three tones still persist in Sango, they function hardly at all to distinguish words or to mark grammatical usage. Were it not for the fact that it is a lingua franca used almost exclusively by Africans, who already speak tone languages, the distinctive function of tone might be altogether lost.

What is lost in the morphology is to some degree compensated for in the syntax. Sango therefore used periphrasis to mark aspect or time. One says fàdé mbi gwe 'I shall go' (literally 'now I go') and mbi gwè àwè 'I went' (literally 'I go is finished').

Lingua francas accompany--in fact, make possible--culture contact and acculturation. It is no wonder therefore that their vocabularies show dramatic changes. This happens when the natural language can no longer compete with other languages, either because it has moved into a different area, or because non-native speakers predominate over the native speakers where the language is indigenous. Thus where Pidgin English is spoken one can usually describe an original situation where learners of English were not only of inferior status but also far more numerous than speakers of English (as in Jamaica, Le Page 1960). The second case might be illustrated by South African Fanagalo or even Town Bemba of Northern Rhodesia.

Lexical borrowing (using words from other languages) and innovations (extending meanings of old words or forming new words with old stock) characterize the aforementioned changes. As an example, witness hay (from 'high') in Cameroon Pidgin English: i mɔni i hay 'his price is high'; i tu hay mɔf fɔɔ tɔk 'He is a loud mouth'; hay stik 'tall tree'; hay skɪn 'proud'; hayam smɔl 'Raise it a little'; ha yu də hayam hayam go bifɔɔ? 'Why are you continually increasing it?'.

When lingua francas are learned under controlled circumstances (e.g. in a classroom) and when there is a recognized value in attaining the normative standard, such linguistic developments occur on a much smaller scale. Yet they are nonetheless revealed in "accepted" dialects of such standard languages as English spoken in West Africa, the Caribbean, or India. Because of their close approximation to the standard there is little doubt that they will be tolerated. Difficulties will arise only when literature produced in the newer dialect is read by speakers of the standard (to some extent this already happens in reading the works of some Indians).

## Fate of lingua francas

What is the future of the languages now spoken by peoples whose cultures are marginal to the prestigious one? With what degree of accuracy can one predict their life expectancy? Assuming an inevitable demise, to what task can they be assigned? Can they accomplish some legitimate goal satisfactorily? These are questions to which responsible government officials and other interested people, such as religious leaders, have addressed themselves. Several fact-finding investigations have already been initiated but their reports are too often filed away with other governmental or private documents. A few conferences have been held. UNESCO convened a meeting of linguists in 1951 to discuss the use of vernacular languages in education (UNESCO 1953). In 1951 there met at Kitwe, Northern Rhodesia, a significant All Africa Conference on Christian Literature and Audio-Visual Communication. In Africa in July 1962 at Brazzaville an international congress of specialists will discuss the subject of multilingualism in Africa. At this meeting due attention will be devoted to the whole subject of lingua francas, and to pidginized languages in particular.

The problem which faces policy makers is fraught with complexities. In every instance they are concerned with welding a heterogeneous multilingual political state into a unified and harmonious nation. They should like to reduce the number of languages. They also want to introduce their countries into the stream of modern civilization. For this they need either a world language in which there already is an important literature or a local language, which must be "modernized." This latter is no mean task. It is difficult enough for Arabic, which had to step from the Middle Ages into the Twentieth Century. How much more difficult it is for an undeveloped lingua franca, such as Swahili.

In many of the emergent nations the official language is actually a marginal language, little known and little used by the population. In Haiti, for example, only about ten per cent of the population can be said to have mastered the language (UNESCO 1951, p. 38); the rest of the population uses

Haitian Creole. Sometimes there is a blindness to the real facts. People in the more prestigious strata have been known to deny the very existence of a pidgin language (like Jamaican Creole) or to claim that "everybody knows" the official language (like French in the Central African Republic). Policy makers must therefore avail themselves of the ever-growing body of literature on all relevant subjects in addition to initiating sociological and linguistic studies; only then can they hope to have established a sound basis for operation.

Linguistic research is necessary to determine the exact nature of the lingua franca. For example, how many different forms of Swahili are there? What are the means in the language to permit it to adapt itself to growing needs? Answers to such questions are being sought for Swahili under the aegis of the East African Swahili Committee (Kampala), but such bodies are rare.

Sociological research must accompany linguistic research to assure a more probable prognosis concerning the outcome of any policy with linguistic ramifications. There are only a few ethnolinguistic monographs in circulation, although several anthropologists are currently interested in related subjects. An inescapable aspect of language use is that it is more than a communication code; it also serves, among other things, to mark ethos identification and prestige.

Some of the most interesting developments are taking place in Africa, which in many respects offers itself as an experimental laboratory to the linguist. Nida (1955) has discussed the linguistic developments on the continent. In a more recent paper by Richardson (1957) (which should be read with Epstein 1959) the competition between several languages as lingua francas is presented. Briefly, there is developing in Northern Rhodesia a form of the Bemba language, a veritable lingua franca, which is gaining ascendance over Nyanja and the two other lingua francas--English and Fanagalo. Five reasons are given: (1) the mine workers in the copperbelt are linguistically heterogeneous; (2) although Fanagalo had been used as a lingua franca until the last war, its identification with the denigrating policies of the Europeans, many of whom used only this language with the Africans, has stigmatized it; (3) the use of English by Africans is opposed by most Europeans, who see in its use an attempt to raise the Africans' status (cf. the discouragement of the learning of German and Dutch in the former colonies of the Cameroon and

Indonesia); (4) Nyanja is identified with the people whose language is the lingua franca of the army and police; (5) Bemba is spoken indigenously by over 60 per cent of the labor force.

## Conclusion

Lingua francas have served tribal communities and vast empires from time immemorial when linguistic diversity prevented social and literary intercourse. Among them have been many drastically reconstituted languages (the pidgins) whose study is currently being undertaken in several quarters. Some dare to say that these are vanishing from the modern world (Whinnom 1956), but the ease with which Korean Bamboo English developed leads one to imagine the inevitability of pidgins in the world.

Historians might feel that all things point to "one world" where internationalism or an international federation or monolithic state would require a single lingua franca. An informed person cannot deny this possibility. But, short of a totalitarian system, the rise of such a language is beset with difficulties.

The possibility of a true world lingua franca would have its advantages. But can one develop without a totalitarian state? Someone has said that Russian is grammatically too complex to serve as such a language; English, however, already meets many of the requirements. In any case, should there ever be an international lingua franca, one thing is almost certain: like all lingua francas before it, it will pass through a stage of pidginization.

## REFERENCES

Bloomfield, Leonard 1933. Language. New York.

Epstein, A. L. 1959. "Linguistic Innovation and Culture on the Copperbelt, Northern Rhodesia." Southwestern Journal of Anthropology 15.235-53.

Greenberg, Joseph H. 1955. Studies in African Linguistic Classification. Branford, Connecticut.

Hall, Robert A. Jr. 1943. Melanesian Pidgin English. Baltimore.

Le Page, Robert B. 1960. Jamaican Creole. London.

Loftman, Beryl B. 1953. Creole Languages of the Caribbean Area. Columbia University M. A. thesis.

Lopes da Silva, Baltasar 1958. O dialecto crioulo em Cabo Verde. Lisbon.

Nida, Eugene A. 1955. "Tribal and Trade Languages." African Studies 14.155-58.

Pierson, Donald 1952. Negroes in Brazil. Chicago.

Richardson, Irvine 1957. Linguistic Survey of the Northern Bantu Borderland: Volume Two. New York.

UNESCO 1951. The Haiti Pilot Project: Phase One 1947-1949. Paris.

UNESCO 1953. The Use of Vernacular Languages in Education. Paris.

Weinrich, Uriel 1953. Languages in Contact. New York.

Whinnom, Keith 1956. Spanish Contact Vernaculars in the Philippine Islands. Hong Kong.

# LANGUAGE SITUATION IN EAST AFRICA

by Ruth E. Sutherlin

## INTRODUCTION

Languages as well as tribes compete for national recognition and hinder the establishment of stable government in Kenya, Tanganyika, and Uganda. Almost 200 African vernaculars, Swahili, and English are spoken by an African population representing almost 200 tribes. If communication and unity rather than isolation of tribal rivalry is to be realized, all these languages can not be emphasized equally in education, government, or the mass media: a common means of communication or the systematic usage of selected languages is prerequisite to effective national government.

This discussion of the language situation in East Africa takes the point of view that language choices made in a tribal context have implications of which educators and administrators should become aware if they would insure success of their programs. This paper considers: I. possible results of unequal language education in a tribal context, II. factors influencing the availability of language education, III. factors influencing the choice of one language rather than another, IV. some choices made by educators and administrators within the specific tribal contexts of Kenya, Tanganyika, and Uganda respectively.

I. The higher one can go in the African education system, the more English one can learn. English is generally introduced late in primary school, made the language of instruction in secondary school, and used extensively in higher education. However, statistics show that relatively few Africans gain secondary or higher education and competency in English. Further, percentages of students enrolled, compared to the actual school age population, show a sharp drop between primary and secondary school.

Those Africans among the linguistic minority who use English well are generally among the political elite. The designation of English as the official language of Kenya, Tanganyika and

Uganda makes it a pre-requisite for active participation in government. Political opportunities are open to English speakers but closed to others. Africans who have not had the opportunity to learn English are dependent on the political elite for translations into the vernacular or Swahili of political concepts derived from English sources.

Because of limited education and the ambiguity or inaccessibility of political information in languages other than English, many Africans find themselves without the linguistic means to participate in national politics. In the swiftly changing political situation which they are not equipped to influence directly, they emphasize tribal traditions and their vernaculars as a source of pride and continuing identity. A tribe which does not want to lose its autonomy to national government can demand recognition for its language, or emphasize its language in tribal or "nationalistic" movements. (In this paper, "nationalistic" will refer to tribal nationalism.) For example, the Gusii in Kenya's North Nyanza District--where Swahili is used in the Legislative Council representing three different tribes with different languages--have requested the formation of a unilingual district. And the Kikuyu of Kenya used their language to promote Kikuyu nationalism during the Mau Mau rebellion credited to them. Tribal identity is a definite link between political and linguistic issues.

Unequal language education often creates dissatisfaction between different generations and sexes, regardless of their tribal identity. For example, most of the older men who would ordinarily hold positions of authority in Kenya are eliminated from direct participation in national government by the requirement that candidates for the Legislative Council pass English tests. Simply because they were too old for public schooling when the British made it available on any wide scale, they did not learn English, and the situation is not corrected by adult literacy campaigns, which are usually conducted in vernaculars. Younger men who do speak English seek the positions in government which their fathers might have held in traditional society, and frequently they must break openly with their families, who regard such strivings as premature or disrespectful.

Women who do not receive education frequently do not speak their husbands' language, even when they move to their husbands' villages or to towns. Whitely notices that Swahili is not spoken by women as much as by men in East Africa; wives of Swahili speakers who have come to towns from their native villages persist in using the vernacular (Whitely, 1956). English speaking Africans are apt to marry English speaking women who can best complement their own political motivations. They recognize inadequate education of women as a problem, however. A Kikuyu saying exemplifies possible dissatisfaction resulting from discrepancy between men's and women's education, "Why learn about sanitation if one must come home and eat from a dirty plate?"

II. The availability of education in general and of language education in particular is influenced (apart from finances available) by educational facilities and their location, number and quality of teachers, availability of text books, library books, and mass media, and policies and attitudes concerning sex and age.

The location of schools influences their curriculum and accessibility. In Tanganyika, Bush schools teach in the vernacular or Swahili, and village schools in Swahili. Middle schools continue with Swahili until Standard VII where English (introduced as a subject in Standard V) is used. Secondary school work is taught in English, and Swahili becomes a subject.

The curriculum of town schools is likely to contain experimental programs which depend on the better facilities, and means for testing available. Children who live outside towns do not have the same general opportunities for higher education or accelerated primary and secondary programs as do urban children. It is common for young Africans to walk miles to school or literally leave home to get an education in a distant place. Travel is almost always pre-requisite to higher education. Makerere College in Kampala, Uganda has been the sole institution of higher general education in all of East Africa.

A shortage of teachers is one of the greatest problems for education in Africa, and the teachers who are available often lack training and consequently produce poor work in their students. There are several programs designed to increase the number of qualified teachers. Americans are being trained by the Peace Corps, and by Makerere College in cooperation with Columbia University. Uganda trains ex-service men as community development teachers and encourages participants in adult literacy campaigns to enlist as teachers for their communities.

One drawback to the use of poorly trained teachers is the potential loss of interest in an experience which lack of training makes unrewarding. An illustrative incident occurred in Uganda in 1959. Primers were prepared in Teso and sold as a pre-requisite for enrollment in a language course. Although 2,500 primers were sold, only 1,181 students could be accommodated, and only 80 were able to finish the course. The high wastage rate was attributed to teachers' lost interest. Even Africans who have the training to make teaching a rewarding experience may lose interest in favor of politics.

Textbooks are available in limited quality and only in selected languages. The content of most of these books is appropriate to an English rather than African context, and may even dull interest in learning. Library books are available in English and Swahili and some vernaculars, but are mostly limited to city populations.

Lacking textbooks, schools turn to the mass media. A number of newspapers in vernaculars, Swahili, and English, and a few periodicals help fill the need for reading matter. Radio broadcasts direct to schools are a common feature of African schools. Cinema vans showing films in Swahili or English accompanied by vernacular broadsheets are used more for adult than public school education.

As indicated above, educational opportunities for adults are limited. Adults are sometimes dependent on their children's better education for keeping up with politics. During the 1961 campaign in Kenya children were seen reading newspapers to their parents. Adult education in mass literacy campaigns is not in English but in the vernacular, a result of the principle underlying teaching in primary school--one should be literate in his own language before one becomes literate in a second.

The need for educated women is recognized in East Africa. After the British manner, separate girls schools are common. The percentage of women enrolled in primary and secondary schools is higher than might be expected from the low overall enrollment.

III. The choice of a language from the many available has political implications. Neglect of a particular vernacular in education can arouse a tribe's nationalism and disrupt an education or government program. In Uganda, a mission committee tried to solve a language problem of a northern province where tribes refused to communicate with one another despite the similarity of their languages: they established a composite language as an official language only to have one tribe whose language was dissimilar demand recognition for it. Other tribes followed suit and the composite language is today known as a "church" language. To guide language choices in education, government, and mass media three indices are suggested: (1) size of a language group and the relation of its language to those of other language groups; (2) political history of the language group; (3) attitudes toward the language and its speakers held by would-be speakers.

> (1) Only three of the almost 200 vernaculars of East Africa are spoken by over 1 million people--Kikuyu in Kenya, Luganda (Ganda) in Uganda, and Sukuma in Tanganyika. Both Kikuyu and Luganda are the languages of tribes with a tribal consciousness that has led them to seek political domination in modern Kenya and Uganda. Foreign administrators have recognized the importance of these languages by printing newspapers in them, providing textbooks, and--in the case of the Baganda--excluding them from education programs involving another language. When tribes resist a program involving a language other than their own, investigation of similarity between the languages involved exposes possible extra-linguistic factors at work. Tribes can claim that other languages are unintelligible simply because they are reluctant to grant another language the same status as their own. The Nandi and Kipsigis of Kenya who speak mutually intelligible languages refused to accept literature written in each other's languages on grounds of unintelligibility. A Nandi-Kipsigis language committee, aware of the similarity of the languages and guessing the political reasons for the claim of unintelligibility, solved the problem by giving the languages equal prominence with the cover term "Kalenjin".
>
> (2) The size of a language group is not an adequate index to its potential demands for attention in education and government policies. The Masai are one of the smallest tribes in Kenya, but they have had an influence in tribal history disproportionate to their size. Reputed and feared as warriors before they were confined to their present reserve, the Masai are the traditional enemy of the Kikuyu. They are known among Europeans

as well as Africans for their pride which keeps them aloof from prolonged contact with other groups. Despite their small size, their poor economy, and their separation by the common border of Kenya and Tanganyika, the Masai try to maintain their tribal unity, and oppose those who would have it otherwise. This persistence suggests that program planners consider the history of language groups which they would include or leave out of programs.

(3) The connotations and attitudes which languages arouse in their speakers and others vary greatly. Vernaculars tend to have nationalistic connotations for native speakers, and threatening connotations for non-native speakers, who fear domination. Swahili, the acknowledged lingua franca of East Africa, is spoken by an estimated 7,000,000 and widely used in both education and the mass media. Yet it is "nobody's language," and does not usually arouse the pride or the fear which the vernacular can. Swahili combines Bantu and Arab elements, borrows heavily from languages of administrators , and reflects interference of other bilingual speakers. There are nineteen dialects --the one of Zanzibar defined as "standard,"--a poetry form, and an abbreviated form developed by European settlers in Kenya's White Highlands.

The history of Swahili's development reveals sources of the varied connotations and attitudes it elicits. In the eleventh century the Shirazi of Persia landed at Pemba off the coast of Tanganyika. They travelled up and down the coast and inland along routes they made famous as the routes of the slave trade. They settled along the way, married Bantu or took them as slaves, and developed for communication in their mixed families and in their business, a language which they named "Swahili," the Arab word for the coasts along which it was first spoken. Swahili is the language of the slave trade, an association which persists today and which has been strong enough to influence the creation of Kingwana, a Belgian Congo dialect supposedly free of Arab elements, which the slaves settling in the Congo as free men wished to eliminate. An Afro-Shirazi party is active in Zanzibar politics. Apparently it arouses the slave trade connotation when politics border on linguistic issues. In response to Zanzibar's current

> insistence that the coastal strip of East Africa should become part of Zanzibar for historical and linguistic reasons, a Kenya party representative replied that they should not make the request because, "it reminds us of the slave trade" (Africa Report, Dec. 16-22, 23-29, 1961).

The Arab elements in Swahili permit it positive connotations as well as negative ones. Swahili has the distinction of being the first African language to be written before the arrival of Europeans. An adapted form of Arabic script permitted the early beginnings of a literary tradition, with its own poetry forms. Organizations for the promotion of Swahili have collected many old manuscripts, transcribed them into roman script, and translated them into English.

Unconcerned with its Arab connotations, the Germans and the British welcomed Swahili as a language of administration, and encouraged its use in education to facilitate their control over Africans. They added connotations of their own to the eclectic language. Among Europeans, Swahili often is reputed to be a language inadequate for the expression of abstract thought or of technical terminology, and thus inappropriate for use in the twentieth century. This attitude towards Swahili may stem from European acquaintance with "KiSettla", the abbreviated form of Swahili originated and used by Europeans for the administration of their plantations in the White Highlands. Described as primarily affirmatives, negatives and commands, it lacks an abstract vocabulary or a technical terminology to an extent that Swahili does not. In Tanganyika Swahili has been proposed as the National language by Julius Nyerere, former Prime Minister and president of the Tanganyika African National Union (TANU), who found it more than adequate for his political campaigning.

In Uganda, Swahili is limited by attitudes toward indigenous languages. It is taught only in the Police Force and in schools for the children of the Police.

Obviously there is no one attitude toward Swahili in East Africa. There is controversy over its continued usage in education. Educators consider dropping it from the curriculum in Kenya, and increasing it in Tanganyika. Although eliminated from Uganda education as a language of instruction, it is admitted to be of surprising usefulness in Kampala, and is recommended to natives who would travel to other parts of East Africa.

Swahili can be a unifying force among potentially antagonistic groups, and in urban contexts where members of diverse tribes intermingle, Swahili has permitted the establishment of new acquaintances and a common life outside the tribal context. In view of high wastage rates in primary school the elimination of Swahili from primary education may seriously deprive those Africans who leave school early and migrate to towns.

Whatever other connotations it may have, Swahili is generally admitted to be "useful".

With English is associated its official position, the political prestige of its speakers, and the increased opportunities it offers students. Africans who have a poor opinion of British administration devalue English: they may prefer French or Russian as a language of wider communication. French is a language which the political elite would like to learn. Despite the obvious benefits of their English education, they are impatient with things English which they associate with delayed independence in Kenya. French offers a means to Pan-Africanism. Russian competes with English and French. There is a Russian-Swahili dictionary already published, and there are opportunities for study in Russia available to Kenya students.

Interviews conducted among African students at an eastern college by Marion Kilson provide a synopsis of language attitudes among multilingual east Africans. Of five students, three preferred their vernacular to either Swahili or English, one for admitted "nationalistic" reasons. All five plan to teach the vernacular, Swahili, and English to their children. Two plan to add French and admit reasons of "Pan-Africanism". Two are themselves studying French, and one wants to study Russian and German. Each language they speak is categorized with a distinct function. English is appropriate for official business, use with foreigners, and politics; Swahili is the language for "other people," "the masses," "other tribes,"; the vernacular is for relatives and "at home." Different subjective feelings are aroused by each language: the vernacular makes one feel "homey," as does Swahili, while English gives one a "formal" or "cosmopolitan" feeling. That Swahili has a positive evaluation is indicated by students' special efforts to speak the language at least three times a week. Students contrast the gains of knowing Swahili and English on personal and social dimensions: Swahili is judged to open avenues of personal communication, "knowing people", and English is judged to open more social opportunities related to status--getting a new job, participating in politics, learning more.

IV. Attitudes toward languages and their use are influenced by the specific tribal contexts, and the history of administration and education policy on language. Some facts for Kenya, Tanganyika and Uganda follow as examples of specific tribal situations and language choices made within them.

KENYA. Kenya has an African population divided among ten main tribes and a number of smaller ones, which make up 16% of the African population. The Kikuyu, who comprise 16.9% of the African population, and the Masai, their traditional rival, have dominated most of Kenya's tribal history. Rivalry is still expressed through the opposition of the Kenya African National Union (KANU) and the Kenya African Democratic Union (KADU), modern political parties for whose origin the tribes are respectively responsible. The Kamba and the Luo, the second and third largest tribes, are joined with the Kikuyu in support of KANU.

The political aspirations of the Kikuyu, Luo, Masai, Kamba and other tribes have influenced language in the mass media. The Kikuyu publish several newspapers with definite political orientations. The Luo and Kamba have four and seven vernacular papers respectively, and the Masai have their own government sponsored paper. Government recognition of tribalism is reflected in the programming of radio networks as well as the use of vernacular in the Press. Regular news bulletins were broadcast in fourteen languages eighteen times a day in the last election, suggesting that, excluding Swahili, English, Arabic, and Gujerati, there are at least ten tribes whose feelings and languages are judged to make a difference in national politics.

Swahili and English both have wide usage in public mass media. There are twenty-seven Swahili newspapers. In the recent campaign, candidates who did not speak Swahili found themselves at a definite disadvantage, and because illiteracy is so high, candidates represented themselves with symbols which could be recognized on ballots. Even Muslim candidates had Swahili speaking representatives on the radio networks. The circulation of English papers is considerably less than that of the Swahili papers and represents a much smaller percentage of African readers than the total circulation figures would indicate.

Education policy reflects the influence of tribalism. Nineteen vernaculars are used in primary school education, Swahili among them along the coastal strip. Swahili is taught as a lingua franca in up-country schools by the third year, and is replaced by English in junior secondary schools. In the coastal towns where Swahili is the vernacular, English is taught from the third year on. Textbooks are available in nineteen vernaculars including Masai.

TANGANYIKA. Tanganyika has 104 tribes listed in the census. They are small in numbers for the most part and often have transitory elements. This extreme tribal diversity and potential instability makes decisive language policy imperative. The acceptance of Julius Nyerere's proposal that Swahili become the national language of Tanganyika would permit greater participation in government for the scattered groups than does English, which is not spoken as widely. Swahili was used by the German administration and continued by the British for both administration and education. Swahili is used in all schools after the first year as the language of instruction until relatively late in the system and then as a subject. The missions in Tanganyika have emphasized vernaculars rather than the lingua franca, however. Forty-four of Tanganyika's newspapers are in Swahili. Only two papers are in Sukuma, the language of the largest tribe, and four in Haya.

Not so much attention is given the vernaculars in either education or mass media as it is in Kenya, suggesting that Tanganyika's tribal situation is one in which assertive nationalistic traditions are not as outstanding.

Legislation to strengthen the position of Swahili among the tribes has been constant in Tanganyika. One of the first groups in East Africa concerned with the standardization of Swahili was the publishing committee established in Tanganyika in 1925. A society for the preservation of Swahili came later, and more recently TANU devoted a cultural study group to the problem of Swahili as a national language. In 1954, regular simultaneous translation in the Legislative Council was introduced. A Swahili Primer has been used for adult literacy campaigns, and Swahili is taught by the government to representatives of both the Peace Corps and the American Friends Service Committee.

The potential usefulness of Swahili to government is illustrated by the Chagga's solution of their linguistic problem--

which Chagga dialect would the government head speak? Swahili was adopted because it was not associated with any group, and hence did not arouse the same nationalistic opposition as dialects.

UGANDA. The Baganda, the largest and most influential tribe in Uganda, have a traditional monarchy that resists plans for representative government that would deprive them of accustomed authority. Other prominent Uganda tribes also form kingdoms of only slightly less strength.

Luganda, the language of the Baganda, has influenced the position of the other vernaculars, of Swahili, and of English in Uganda. In 1930 when the Inter-territorial Committee on Swahili was formed, Uganda agreed with Tanganyika and Kenya on the importance of Swahili for education. But three years earlier the Kabaka (king of the Baganda) had issued a public statement protesting the introduction of Swahili into the school system, on the grounds that it threatened loss of tribal status which the Kabaka was dedicated to protect. (The government responded with an explanation that they never planned to include the Baganda in the program.) Missions in Uganda objected to Swahili because of its Arabic words and their associations with the Muslim religion.

Today the position of Swahili in Uganda has declined considerably. Swahili is taught nowhere but in the schools for police and their children, a silent admission of Swahili's usefulness as a means of inter-tribal communication necessary for police work. The very fact that Swahili is "nobody's language" (the Swahili population of East Africa, centered in Tanganyika and Zanzibar is not a tribe, but isolated individuals born of Arab and Bantu parents) caused a reaction of distaste to it among Uganda's powerful kingdoms.

By 1944 a Committee for the Standardization of Ganda was formed as part of a movement to increase Luganda's use in Uganda. So emotionally involved were the Baganda with the decisions of the committee, that two years after the standardization an African editor was beaten for using the "wrong" form. (The Kabaka had formed an African Editors Club in order to prevent "wrong" language usage in the press, and had assumed powers to demand correction of such usage when it occurred.)

Despite the influence of the Kabaka and the Uganda, Luganda did not become the national language. In part, the power of the Baganda prevented the dominance of their language: Baganda served as administrative agents for the British and aroused other tribes' hostility against themselves and Luganda. Six literature committees for languages other than Luganda were established to promote languages other than Luganda, particularly in education. In 1947, Luganda, four other vernaculars, and Swahili were educational media. Demands were made for introduction of the two other languages represented by the committees. At present eight vernaculars are used in primary school during the first three years. More and more stress on English occurs. Teachers frequently introduce it before the prescribed middle standard level. In anticipation of increased emphasis of English in education, the Institute of Education at Makerere is investigating the problems of transition from the vernacular, especially Luganda, to English. Textbooks in languages other than English are available in Luganda, Nyoro, Lwo, and Swahili, but in limited supply. Though Luganda does not dominate the schools, it does the press. There are fourteen Luganda newspapers, five in another vernacular, three in each of two others, and one in each of three others. The Baganda's past resistance to language programs that would be introduced or extended at the expense of Luganda predicts their resistence to efforts to elevate English over Luganda in modern Uganda.

## BIBLIOGRAPHY

Allen, J. W. "The Collection of Swahili Literature and its Relation to Oral Tradition and History," Tanganyika Notes and Records, 53:224-228, 1959.

-----. "The Rapid Spread of Swahili," Swahili, Journal of the East African Swahili Committee, 30:70-74, December 1959.

Baker, E. C. "Notes on the Shirazi of East Africa," Tanganyika Notes and Records, 11:1-10, April 1941.

Broomfield, G. W. "Development of the Swahili Language," Africa, Vol. III, 4:16-22, 1930.

-----. "Re-Bantuization of the Swahili Language," Africa, Vol. IV, 1:77-85, 1931.

Gower, R. H. "Swahili Borrowings from English," Africa, Vol. XXII, 2:154-156, 1952; reprint, Tanganyika Notes and Records, 50: 118-120, 1958.

-----. "Swahili Slang," Swahili, Journal of the East African Swahili Committee, 28:41-48, July 1958.

Harries, L. "Congo Swahili," Tanganyika Notes and Records, 44:50-54, September 1956.

-----. "Swahili Epic Literature," Tanganyika Notes and Records, 30:73-78, January-June 1958.

Kilson, M. "Multilingualism among African Students," (unpublished manuscript.) Harvard, 1962.

Reusch, R. "How the Swahili People and Language Came into Existence," Tanganyika Notes and Records, 34:20-27, January 1953.

Robinson, A. E. "The Shirazi Colonization of East Africa," Tanganyika Notes and Records, 3:40-81, 1937.

Roehl, K. "The Linguistic Situation in East Africa," Africa, 3:191-202, April 1930.

Sutherlin, R. "Interdisciplinary Approach to Tribalism," (unpublished manuscript.) Stanford, 1961.

UNESCO. The Use of Vernacular in Education (Monograph on Fundamental Education.) Paris: UNESCO, 1951.

-----. African Languages and English in Education (Education Serials and Documents.) Paris: UNESCO, 1953.

-----. The Education Situation in Africa Today; Final Report of the Conference of African States on the Development of Education in Africa. Addis Abbaba: 1961.

U. S. National Commission for Unesco. Africa and the United States; Images and Realities, 8th National Conference, (Background Book) Boston: 1961.

Whitely, W. H. "The Changing Position of Swahili in East Africa", Africa, Vol. XXVI, 4:343-354, October 1956.

-----. "Language and Politics in East Africa," Tanganyika Notes and Records, 47-48:159-174, June and September, 1957.

-----. "Swahili and the Classical Tradition," Tanganyika Notes and Records, 53:214-224, 1959.

-----. "Political Concepts and Connotations, Observations on the Use of Some Political Terms in Swahili," in K. Kirkwood (ed.) St. Antony Papers No. 10, African Affairs: Number 1. Carbondale: Southern Illinois University Press 1961. Pp. 7-22.

Whitely, W. H. and Gutkind, A. E. Linguistic Bibliography of East Africa, Revised 1958. Arusha: Beauchamp Printing Company for the East Africa Swahili Committee and the East African Institute of Social Research.

# LANGUAGE PROBLEMS IN THE RURAL DEVELOPMENT OF NORTH INDIA

by John J. Gumperz

Most of the discussion regarding the language issue in India so far has revolved about two problems. One of these concerns the spread of Hindi knowledge in those areas or fields of endeavor where other regional languages or English are now used; the other, the enrichment of Hindi vocabulary (or, as is usually stated, the development of the Hindi language), so as to make it possible to use it for dealing with topics related to modern science and technology. The Central Government and several of the state governments have devoted much effort towards the solution of these problems by setting up special departments for the creation of vocabulary lists to deal with new subjects, encouraging attempts to introduce Hindi in universities and legislative bodies, and fostering the growth of societies for the propagation of Hindi learning. These activities have aroused a great deal of controversy, and so much has been written about the pros and cons of the issues involved that the above two problems seem to be the only ones associated with language in the public mind.

There is, however, another language problem of a slightly different nature which has received little attention so far. This concerns the difference between the everyday spoken language of the people and literary language used on formal occasions and especially in writing. It applies to all parts of India, including even those areas in which Hindi is now the regional language. It is common knowledge that in addition to Hindi and the various regional languages recognized by the Indian Constitution, a number of dialects are spoken on the village level, but very little is known to nonlinguists about the difference between these dialects and the

---

From the Journal of Asian Studies 16.251-59 (1957). Reprinted by permission of the author and the Association for Asian Studies, Inc.

standard spoken and written language and the extent to which the three are mutually intelligible.

The present paper is an attempt to illustrate this problem on the basis of observations made during eighteen months of linguistic research in various parts of rural North India.[1] The area chosen for illustration is the area in which Hindi is now the regional language. The problems discussed, however, apply to a greater or lesser degree to all parts of the country. The paper consists of two parts. Part one is a short descriptive statement of some of the facts of geographic and social language distribution in the so-called Hindi-speaking area; the other is an account of some of the linguistic difficulties in communication that have come up between extension workers and government officials and villagers in a North Indian village.

The gap between popular speech and written language has been in existence during the greater part of Indian history.[2] Sanskrit had long ceased to be a spoken language in the times of Kalidasa, the greatest of its writers, and even the Prakrits, which are used by lower-class characters in Sanskrit plays, do not seem to have been accurate reflections of popular speech. In the present discussion, an attempt will be made to draw a clear distinction between literary or written idiom and spoken language.

Before going into the matter of language distribution, an attempt will be made to explain the various uses of the terms Hindi and Urdu, regarding which there seems to exist a great deal of confusion in non-linguistic literature. The word Urdu was originally used to designate the literary idiom which had gained currency in the Dravidian-speaking country of Golconda on the Deccan Peninsula in the late sixteenth century and which was based on the spoken lingua franca used around the Moghul army camps and administrative centers.[3] This language in turn had had its roots in the local dialects spoken in the Gangetic Doab, that is, the area around Delhi and Merut. Urdu literature was written in an adaptation of the Persian script and had taken over a great number of words from Persian. During the eighteenth century it spread to such North Indian urban centers as Delhi and Lucknow. In the beginning of the nineteenth century, a number of Hindi writers began using the Devanagari alphabet for their writings and started replacing many of the Persian borrowings with loan words taken from Sanskrit. The new idiom which thus developed was given the name of Hindi. Literary Hindi and Urdu are thus two styles of the same language; they are used in writing and, on a spoken level, for formal lectures and discussion. Both have the same inflectional system and a common core of basic vocabulary; they differ in the learned or abstract words used and in word order. Spoken literary Urdu also has taken over a number of sounds from Persian, such as those represented by *q* and *g͡h* which have fallen together with those of *k* and *g* in spoken Hindi.[4] The sounds equivalent to *kh*, *z*, and *f*, which also come from Persian, are commonly used in both styles. Literary Hindi, on the other hand, uses a number of initial and final consonant clusters which are not common to Urdu.

Side by side with the above two literary idioms, the spoken lingua franca from which they had sprung continued to spread in the bazaar towns and urban administrative centers. Its spread was favored by the fact that it had become the official medium of the British administration and also by the great improvement in communication facilities during the nineteenth century. Outside the area where Hindi is the regional language, it is now widely used in such urban centers as Karachi, Bombay, Calcutta, Hyderabad, and others. In British India it had usually been referred to by the name Hindustani. After Indian Independence this latter term was dropped and the word Hindi was adopted for the national language of India. This change in terms, however, has not resulted in a corresponding change in language.

In spite of the fact that there is a large relatively active group of people who favor the elimination of many colloquial words in favor of Sanskrit borrowings, the gap between the spoken language (hereafter referred to as standard spoken Hindi) and the literary idiom remains. This fact is usually ignored in the non-linguistic literature on language problems. The confusion in terminology is further increased by the popular use of the terms Hindi and Urdu. Hindus of every class are now apt to refer to their own speech as Hindi, whether they speak a village dialect or the standard language. Muslims, on the other hand, refer to their speech as Urdu, although from the point of view of the hearer there might be little difference between their speech and that of their Hindu neighbors. The author has had several experiences in which Hindus in complimenting him on his efforts to speak the local idiom would say, "You speak Hindi well," whereas Muslims would react to the same or similar words with, "You speak

Urdu well." A formal analysis of at least one village dialect has shown that the differences between the speech of touchable and untouchable castes are much greater than those between that of Muslims and Hindus.[5]

The area in which Hindi is the regional language extends from Rajastan and the eastern part of Punjab in the west to the eastern borders of Bihar and Madya Pradesh. The standard spoken Hindi, which is current in the urban areas of this region, is relatively uniform. The literary Hindi idiom is taught in the schools. The introduction of Hindi literature has displaced a number of older literary idioms in which there had been literatures since the late Middle Ages. The best known of these are Braj Bhaśa, Avadhi, and Maithili. At the present time literary activity in these idioms has died out, but the spoken forms on which this activity was based continue to be used by educated as well as uneducated city residents in the home and when among friends. Most city speakers of these dialects are equally at home in Hindi, however, and actual conversations are characterized by frequent switching from the dialect to Hindi and vice versa. These dialects occupy a somewhat intermediate stage between standard spoken Hindi and the local village speech, since they are spoken in relatively uniform form over a large area. They will be referred to hereafter as regional standards.[6] In addition to those mentioned above, there are a number of regional standards which have had no literature, e.g., Marwari and Jaipuri spoken in Rajastan, Surguja in Madya Pradesh, Magadhi in Bihar. Some of these regional standards differ only slightly from Hindi; others must be classified as different languages from a linguistic point of view. Thus for example, Braj Bhaśa, the language spoken around Agra, differs little from standard Hindi. The Jaipuri and Marwari dialects of Rajastan, on the other hand, show many more differences. They have a much greater variety of inflectional noun and verb forms. The differences between their inflectional system and that of Hindi are so great that linguists generally classify them as part of a separate language, Rajastani, which as Grierson states, is more different from standard Hindi than Punjabi, one of the regional languages recognized by the Indian Constitution. The greatest amount of linguistic difference within the Hindi-speaking area is found between Maithili and Hindi. Maithili has a vowel system which is akin to that of Bengali and a system of verbal inflection in which the verb is inflected for person and according to the social status of both the subject and the object, and not according to person and number as in Hindi.[7]

On the level of village speech, the amount of linguistic variety is even greater. The saying goes in rural India that each village has its own dialect. This is of course an exaggeration. In the Gangetic Doab, however, the region where standard Hindi first developed, differences are so great that villagers from Moradabad District state that they have difficulty in understanding the speech of their own relatives by marriage from Karnal District, slightly more than a hundred miles away. These local dialects have been studied very little by linguists so far. Grierson's survey is based on written answers to questionnaires administered through British official channels and not on direct field work; most of the other publications on Indian dialects deal with regional standards.[8]

From a linguistic point of view the local dialects form a continuous chain from Sind to Assam, the speech of each area shading off into that of the adjoining one. The individual speech differences can be plotted on a map with lines showing the geographical boundaries of usages, but in no two areas adjacent to one or a group of such boundaries are the differences so great that there is no mutual intelligibility. Lack of mutual intelligibility is found only between areas that are relatively far apart. The regional languages such as Sindhi, Gujerati, Marathi, and Hindi, and also the abovementioned regional standards have been, so to speak, superimposed on this chain of dialects as a result of political and historical factors. If therefore we take two villages, say on the borderline between Rajastan and Gujerat, we will find that villagers will have no difficulty in understanding each other, whereas government officials who speak only the standard languages might have to resort to interpreters in order to communicate effectively. If the official from Rajastan knows the regional standard, communication will be easier. Learning a literary idiom for a villager, on the other hand, often requires much more than mere acquisition of literacy--it may be equivalent to learning a foreign language.

The greatest amount of conservatism and inflexibility with respect to local dialects is found among those groups in the village that have little opportunity for contact with outsiders, such as women and certain members of the lower castes. Outside the village, when talking to traders and officials in the bazaar towns and administrative centers, villagers are forced to drop some of their most divergent localisms and adapt their speech to the regional standard. It is also in centers of this type that standard Hindi is becoming more and more current. Many of them contain settlements of traders from different areas, e.g., Marvaris from Rajastan, Sikhs from Punjab, Muslims from Lucknow or Delhi, who use Hindi to communicate with the local population. Other factors which aid in the spread of Hindi are the improvement of communications, which has made long travels rather painless, the schools and the movies, which may be found in the smallest of towns. Data from one village in the Bhojpuri-speaking area of eastern U. P. shows that a large proportion of the local men understand and speak some Hindi; the women and many of the low-caste people, on the other hand, have trouble in following it. It must be remembered, however, that the language which is spreading is the spoken standard Hindi and not the literary idiom. Except in the schools, the average villager has very little opportunity for coming in contact with the literary language, since the amount of printed material that reaches him is still very small. It is also true that the conversations in standard Hindi that the ordinary villager has occasion to engage in revolve about such everyday matters as common greetings, commercial transaction, etc.; conversations about philosophical problems or the more intimate problems of family life and religion would be more difficult to carry out in Hindi. Anthropologists working in the Gangetic Doab have found difficulties in discussing such matters without resorting to the local dialect.

The studies of written communication were made in a village in Saharanpur District of Uttar Pradesh. The dialect of the area is referred to by Grierson as Vernacular Hindustani and it is one of the group of local dialects which show closest affinity to standard Hindi. The standard speech is universally understood, although the village speech itself is not easily understandable to outsiders from different regions.

The village is located in a fertile agricultural area, which has in recent years become one of the centers of Indian sugar-cane production. Cane is the principal village crop. There are two large sugar mills at a few miles distance, and between the months of November and May, during the sugar-cane season, cultivators make several trips a week to the mill to deliver cane. The tahsil (sub-district) headquarters, a town with a population of about 25,000, is located about six miles away on the main railroad from Delhi to Saharanpur. Local roads are quite good and rail and bus connections with Delhi to the south and Ambala and the Punjab to the north and west are excellent. Villagers have marriage relations as far away as

Ambala in the Punjab and Bulanshaher and Moradabad towards the east. Many of them have had several occasions during their lifetime of travel within this entire area in connection with family or business affairs. The primary school in the village dates back to the 1920's. An inter-college (equivalent to a senior high school) has been in existence for seven years. It is supported primarily by village funds and many of the local boys attend. The C.D.P. (Community Development Project) has been active in the village since 1953. All in all, the village seems exceptionally well provided, as Indian villages go, with educational facilities and opportunities for contact with the outside world.

From the point of view of literacy, the adult population, i.e. people between the ages of twenty-five and fifty who carry the major responsibility of village life, can be divided into three groups. The literates are those that can read well and regularly read newspapers, novels, or other literature. The majority of them have had five or more years of schooling. Their first instruction was in Urdu, since until Independence Urdu was taught in the primary schools of the area. They have had to learn the Hindi script and literary idiom for themselves. There is another group of semi-literates which includes those who at one time or other have had some schooling but read with difficulty, or those who have taught themselves to recognize the letters of the alphabet. Many people in this group have had up to five years of schooling; some of them say that at one time they read well but have forgotten now. This group is numerically much larger than the first. It is constantly growing since many people have begun to learn to read, either on their own or with the help of neighbors. The last group, that of the illiterates, includes those that do not have a knowledge of the alphabet and is approximately equal in numbers to the other two groups combined.

The intelligibility tests were conducted with a pamphlet, distributed by the C.D.P., announcing a development fair. This pamphlet contains forty-five lines of text announcing the various features of the fair. It was tested with a panel of informants drawn from the three above-mentioned groups of literates, semi-literates, and illiterates. The methods used for testing were patterned on the traditional way in which written materials are made known to illiterate villagers. This is usually done by reading the pamphlet out loud to the audience, paraphrasing the text in the village dialect and commenting on it. In the present tests the text was read to the illiterates; semi-literates and literates read the text themselves; all informants were asked to explain the meaning as if they were telling somebody else about the pamphlet.

The illiterates were able to obtain only the vaguest notion of the contents of the material read to them. They made mistakes in understanding in ten to fifteen of the forty-five lines of text. The semi-literates did slightly better; they averaged about ten mistakes. The literates understood the main part of the message; however, even they did not know some of the terms used. In passages which dealt with something familiar, such as those features that are usually associated with a fair in the experience of villagers (e.g. wrestling contests, shows, cinema), the informants were able to guess at the contents of a passage from a few words. The purpose of the fair, however, was to stimulate interest in new ideas, such as crop competitions, public health, domestic industries, the participation of women in community life. The greatest number of misunderstandings occurred in the passages relating to these last features.

From a linguistic point of view the difficulties encountered were largely in the realm of vocabulary. Here are some examples. The announcement referring to a baby show read: *tīn sāl tak kī āyū kē chōṭe bālkō̃ kē svasty kī pratiyōgitā hōgī,* 'there will be a show of healthy babies up to the age of three.' This was interpreted by several people as: 'There will be a wrestling match of three-year-old children.' The difficulty lies in the use of the terms *pratiyōgitā* 'show' and *svasty* 'health' instead of the common words *numāysh* and *hālat*. Another example of the type of misunderstanding that occurred is the following: *mahilāõ̃ kē līyē ēk din vishēś prakār sē rakkhā jāyēgā* 'a special day will be set aside for women.' The word for women *mahilā* is somewhat similar in phonetic shape to the local form *mahōllā* 'neighborhood,' and the words *vishēś* 'special' and *prakār* 'manner' were not understood. The common interpretation of the passage was, 'people from every neighborhood are invited.' Only the literates understood the passage. Examples of other words that were not understood are: *virāṭ* 'big', *ev* 'and,' *avsar* 'occasion,' *unnatī* 'improved,' *kṛsi* 'agriculture,' *gṛh udyōg* 'domestic industry,' *lōk gīt* 'folk song,' *sammēlan* 'assembly.' Many of these indicate exactly those concepts that the C.D.P. has been attempting to introduce. It was rather surprising to see that the official name of the C.D.P. center, *samudāyik vikās yōjnā kēndr*, which was placed above the title of the pamphlet, was known to only one among all the informants interviewed. The common word in the village is the word *parōjak* 'project.' The village-level worker is usually referred to by the term *dabalū* 'double u' (from V.L.W.) or *naspattar* 'inspector' (the name given to agricultural officials during the British regime), instead of the official term *grām sēvak*.

Some of the other difficulties in understanding were due to the use of words in a meaning different from the meaning they commonly have in the village; e.g. the use of the word *dhangal* to mean 'contest' in one case caused great difficulty for several informants who knew only the village meaning 'wrestling match.' Difficulties in syntax were rare; a few informants had some trouble with word order, but they were able to make it out after going over the pertinent passages a few times. It should be expected that syntactic difficulties would be greater with speakers of languages more different from Hindi such as Rajastani or Maithili.

Difficulties in comprehension similar to those mentioned above were found to a greater or lesser degree in most of the literature that reaches the village, such as in farm magazines, newspapers, and the few extension leaflets that are passed out. In one leaflet, which was distributed by the local school authorities on the occasion of a visit to the village by the President of India to lay the foundation stone of a hospital, the words for hospital and foundation stone were not known to any of the informants. Illiterate and semi-literate informants also had great difficulty in understanding an article on better methods of cotton cultivation in an agricultural magazine which regularly reaches the village.

Lectures on technical subjects by outside officials are also likely to present difficulty. The village level worker and the officials from the local development center who have been in the area for awhile have no trouble in making themselves understood. There have been some lectures, however, by outside technical experts, notably one on artificial insemination, which were understood by only a few of the literate villagers. There are five or six radios in the village, but people do not ordinarily listen to the All India Radio News broadcasts because they say the language is too difficult for them.

In the relations of villagers with traders from the bazaar towns and with the courts, another type of difficulty arises. Urdu is still used by many businessmen for keeping accounts and for posting written notices. The account sheets issued by the sugar mills, listing the deliveries of cane, are printed in Hindi, but the local clerks often write in the figures in Urdu. Urdu is also still used in some of the land records. There are, furthermore, a few cases where instructions in English have been issued to the sarpanch, the head of the village tribunal, by the superior court official. The person in question happens to be literate in both Urdu and Hindi, but since he knows little or no English, the instructions have so far been disregarded. In some cases persons involved in a trial are furnished only English transcripts of the judgment in their case. There are two railroad crossings near the village, on dirt roads travelled primarily by bullock carts, where the only warning signs are in English.

If a villager therefore wants to be able to read the sort of literature that reaches him in the village and thus remain in direct touch with the outside community, it is not enough that he know the Devanagari alphabet. In order to read Hindi he has to learn a great deal of new terminology and new syntactic constructions. In order to transact business he has to master the Urdu script and Urdu literary style and it would also be useful if he knew at least some English. In the absence of this knowledge he is forced to depend on middlemen, whose reliability he has no way of judging.

From the point of view of rural development, the local language situation imposes severe limitations on the methods that may be used for spreading new ideas. The only reliable means of communicating with villagers is personal contact by the village level worker or a local project officer. Written communications at best reach only the relatively small group of literate villagers, and the probability of misinterpretation by readers who are unable to comprehend the contents and pass on their misapprehensions to their illiterate neighbors are quite high. The present language situation further serves to preserve the traditional barrier between the villager and the outside world, a barrier which the C.D.P. is trying to remove.

The effectiveness of development might be greatly increased if the diversity in style and language used by local government agencies could be eliminated and all persons concerned could be trained to use a style intelligible to villagers, at least in those writings that are destined for village consumption. It would then be possible to reach the ever increasing group of semi-literates, many of whom might otherwise revert back to illiteracy. There have been some attempts made to train writers in the production of literature for new literates, but the writings produced have so far not reached the village in which this study was made.[9] In any case, what is needed is a co-ordinated effort on the part of all agencies concerned, not just the distribution of a few booklets.

The data presented in the foregoing is based on only one village, but the impressions gained have been confirmed by other researchers with experience in Indian villages.[10] Furthermore, the linguistic diversity described in the first part of the paper seems to indicate that similar conditions exist also in other areas. The problem seems real enough to suggest further study, based on a more detailed regional survey of language problems at the village level.

## REFERENCES

1. The field observations resulting in this study were made under a fellowship granted by The Ford Foundation. The conclusions, opinions, and other statements in this publication, however, are those of the author and not necessarily those of The Ford Foundation. The author is indebted to Professor Morris Opler and the staff of the Cornell University India Project for furnishing living quarters in the village and providing much helpful background information for the study. Special thanks are due to D. S. C. Dube for his help in formulating the problem.

2. India is of course not unique in this respect. Similar differences existed throughout medieval Europe and still exist, for example, in Greece, where they have become the subject of a great deal of political controversy.

3. Sir George Abraham Grierson, Linguistic Survey of India (Calcutta: 1927). For an account of the rise of spoken and literary Hindi see Vol. I. Another account of the development of modern Indo-Aryan vernaculars, which also discusses the gap between spoken language and literary idiom, is given in S. K. Chatterji, Indo-Aryan and Hindi (Ahmedabad: Gujerat Vernacular Society, 1942).

4. The Library of Congress transcription of Devanagari as adapted to Hindi is used throughout.

5. John J. Gumperz, "Dialect Difference and Social Stratification in a North Indian Village", American Anthropologist 60.668-82 (1958).

6. Grierson does not draw this distinction explicitly; however, it is implicit in his comments.

7. See Grierson, IX, Pt. 2, p. 1 for Rajastani, V, Pt. 2, p. 13 for Maithili.

8. Some excellent examples of studies of this type are: Baburan Saxena, Evolution of Avadhi (Allahabad: 1937); Chirendra Varma, La langue Braj (Paris: Librairie d'Amérique et d'Orient, 1935).

9. Workshops on writing for new literates were held in Delhi in 1953 and in Mysore, Poona, and West Bengal. The Indian Council of Agricultural Research publishes a journal, Dharti ke lal, intended for new literates, which has a circulation of about 8,000; cf. The Ford Foundation and Foundation Supported Activities in India (New York: The Ford Foundation, 1955). The Uttar Pradesh Development Commissioner at Lucknow has also published a series of pamphlets on agricultural topics for new literates.

10. Mr. Philip Barker, a linguist who has spent about eighteen months in rural areas of Sarguja, Madya Pradesh, tells me that villagers in his area were not able to understand the terminology of official announcements posted in the village. See also Chatterji, pp. 140f.

# LANGUAGE STANDARDIZATION

by Punya Sloka Ray

## Introduction

We ordinarily speak of standardization in relation to tools. We expect of a standardized tool that it will be cheaper to acquire and maintain, that individual specimens will be very much alike and of relatively uniform dependability. When a tool is rarely used or used by only a few people, standardization is relatively unimportant. But if it comes to be used frequently and by a large number of people, standardization is often an advantage.

When we apply the concept of standardization to languages, we stress their tool-like character. From this point of view a language is only an instrument of communication; a means, not an end. In the following discussion we consider questions of efficiency and uniformity of linguistic practices, as well as questions of policy in furthering standardization, especially insofar as it can be a deliberate and conscious operation.

In this discussion we confine our attention primarily to a single aspect of language: lexical forms, i.e. the vocabulary (as distinguished from the grammar).

## Questions of efficiency

**The first thing is to dispel the notion that all linguistic practices are equal, that there are no real choices available on grounds of relative efficiency. However, a preliminary point is that we cannot simply set up an abstract scale of efficiency and grade each word or phrase of the language upon it. The expression which may be less efficient for certain purposes may be more efficient for others. We need to have therefore both an assessment of identifiable purposes and a canon for choice when an identified purpose and a small set of alternative forms are given. In brief, we do not require a method for measuring efficiency, but only for comparing it.**

Let us consider what happens when the expression "vegetables" is replaced by "veg". This is often justified on the ground that the former is a mouthful compared to the latter. Yet it is also clear that "veg" requires more explanation than "vegetables". If and when the unabbreviated form is known and is not too rare, and the general technique of the abbreviation, such as 'pronounce only the stressed syllable', is even better known, then the explanation required for the longer form may not counterbalance the advantages gained by the shorter form. Only in such a case may the shorter form be said to be more efficient. We shall call this type of efficiency 'concision'.

But there is another aspect to efficiency. Suppose that I am talking to a friend over the phone and he is being distracted by other things. In such a case, he may easily miss "veg" and ask me to repeat. "Vegetables" would not be so easy to miss. For example, if he misses "v.g...b...", but catches ".e.eta.les", there are not many alternative expressions that would fit, given the general context of the conversation. This quality has been called 'redundancy' but may be also called self-confirmation of the utterance.

We shall also need to distinguish between two different concepts of frequency. 'Text frequency' compares two lexical forms in their repetitions with a body of discourse. For example, in a text of 5000 words we may find that the word "set" has been used 45 times while the word "form" has been used 63 times. 'List frequency' compares two lexical forms in their repetitions of pairing with other lexical forms. For example, we may find that within the same text the word "set" has been used immediately before 35 different words while the word "form" has been used before 39 different words.

A lexical form may also be identified as either relatively 'familiar' or relatively 'learned'. It is more familiar when it has a higher list frequency and more learned when it does not. The more familiar a lexical form is, the more efficient it is to have it name a relatively fluid concept; the more learned, the more efficient it is to have it name a relatively precise concept.

Let us consider the role of associations. The higher the list frequency of a lexical form, the more diverse associations it has and if we use a more familiar word to designate a precise concept, we run the risk of being led off in irrelevant directions. But because the learned form has a lesser burden of associations, it can be used more arbitrarily and precisely than the more familiar form.

## Questions of uniformity

There are two different ways in which uniformity may be sought for a language. First, we may have a set of people who as a group identify themselves as users of the language, and yet who do not use the same forms for the same meanings. This situation is usually described as existence of dialects. All languages are fragmented into dialects, but not to an equal degree. For instance,German may be said to be more fragmented into dialects than French. That the causality of such a phenomenon might be found in another dimension, that dialects might correlate with socio-economic disparities in the population, is not relevant for the point of view adopted here. What matters here is only the ability to contrast languages. If the forms and meanings vary relatively little between smaller groups within the larger group of the speakers of the language then the language is relatively 'closed'.

We may have a set of people who as a group identify themselves as users of the language, and yet who do not use it for all purposes of life. For instance, up to about the second world war, many of the chemistry textbooks in use in American universities were in German, and researchers in that subject were expected to be able to conduct discussion in that language. Indeed, not being required to use a language other than one's vernacular for some urgent purpose or other is a rather exceptional situation in history. This phenomenon leads to the question how far the forms and meanings of the first language match with those of the other languages also needed. This we shall describe as the relative 'opening' of the language. We might speak, instead of 'closure' and 'opening', of intra-linguistic and inter-linguistic uniformity.

In summary, lack of uniformity may be shown in various ways in particular cases. [Portions of the following treatment through the end have been published as part of an article in QUEST, October 1961, the courtesy of the editor of which is hereby acknowledged.] Firstly, there may be no common, uniform, standard linguistic practice at all within a society. Secondly, there may be one common, uniform but functionally restricted standard. Thirdly, there may be several functionally restricted standards.

Source of borrowing: 'Borrowing' is a traditional term for a phenomenon we might better describe as 'inheritance': no contract for repayment is involved. By a source of borrowing is meant here a source which is currently utilized, and not sources that may have been utilized in the past. English did borrow heavily from Danish at one time. But Danish is not a source of borrowing for English now, at least not a major source.

These sources may form a serial order functioning almost as one single whole. This is the case with English, for which the major sources are French, Latin and Greek. The point to note here is that French borrows heavily from Latin and Greek, and that Latin used to borrow heavily from Greek.

Or the sources may be ordered parallel to one another. This kind of an order is exhibited for Indonesian, which borrows mostly from Dutch, French, and English. Similar is the case for German, for when it borrows it usually does it from either French or English.

Or the sources may be ordered in two unrelated and divergent groups. This is the case with Japanese, which borrows from classical Chinese as well as from modern European languages. The situation is similar for any modern Indian language, with Sanskrit (or classical Tamil or Arabic-Persian) and English as the sources. The tension which exists in the last type of relationship may be alleviated in various ways. First, one of the sources may be abandoned. This has been done in Turkey. Many words traceable to Arabic or Persian are still used, but fresh borrowing from Arabic or Persian has stopped. Second, elements from one of the sources may be used in ways which harmonize with the practice of another source rather better than they do with the original usage. Many of the new formations which German produces by recombining native elements, or Japanese by recombining elements borrowed from classical Chinese, are more or less mechanical translations of the Latin-Greek terms current in the modern European languages. Third, elements and combinations from both directions may be accommodated as stylistic variants. Japanese practices this method to the extent of matching Chinese-derived script-forms and European-derived sound forms for the same Japanese-assimilated words. One might reproduce a sample in a hypothetical English by spelling "artificial-satellite" but pronouncing "sputnik".

Dialect: There may be within a single language no one dominant dialect which can serve as the base for a common standard. This was the case in Germany before Luther, with about five major dialects. However, Luther's choice, combined with the otherwise admitted advantages of the relatively most unified dialect, did tip the scales towards a clear decision. Frequently, there is no clear awareness of the situation in this regard. This was the case in Southern Rhodesia, where Clement Duke had to research for a year before being able to recommend in favor of the very closely related Karanga and Zezuru dialects for building a written, unified Shona.

Or there may be a common, uniform but functionally restricted standard dialect, limited to formal oratory and literature. This phenomenon has been termed 'diglossia' and is well illustrated by Arabic. Formal Arabic differs little from the classical language and is used all over the Arab world. But ordinary conversation even among the highly educated is carried out in several distinct dialects in Iraq, Syria, Egypt, and the Maghreb. The long range outlook, given present political disunity, seems to favor the rise of several daughter languages on the model of the rise of French, Spanish, Italian etc. from Latin. A second possibility is the eventual dominance of one of the spoken varieties over the rest. This seems to be the drift in mainland China, where the Peking dialect is being pushed to replace all others. A third possibility is the indefinite persistence of the situation. This is the case in Switzerland's German-speaking areas, where High German and the Swiss German dialect supplement each other.

Or there may be more than one common and inadequate formal standard within the same language. One example is the situation for some decades and in some areas of Bengal. There were two formal standards, one without a spoken base in the present and the other with a spoken base in another region, (besides several fairly stabilized spoken dialects structurally intermediate to the two). The situation has now been resolved by a rapid obsolescence of the first of the literary dialects.

Language: There may not be a single common language. This is rarer these days than might be supposed. Two examples are Finland and Switzerland. In Finland, Finnish is spoken by over 90 per cent and Swedish by less than 9 per cent. But Swedish was until late last century the official language, and still is the nearest equivalent to a lingua franca in the Scandinavian region. In consequence, Swedish is the official second language and the urban population is also actually more or less bilingual. In Switzerland, German is spoken by about 72 per cent, French by a little over 20 per cent, and Italian by slightly less than 6 per cent. But except for formal purposes the German speakers use a rather pronounced dialect, while French and Italian are spoken much as in France and Italy, which are powerful and often admired neighbors. Moreover, Switzerland was for a long and crucial period of her history under the influence of France. The net result is that the language most often utilized as a second language within the country is not German, but French.

Or there may be one common but not quite adequate language. In the Philippines, English is the common language and, from the third grade onwards, the only medium of instruction. The official census declares with perhaps some exaggeration that 38 per cent of the people are actually able to use it. English is clearly in the process of being transformed from an originally foreign, then upper crust second language to a mass language, although the native languages are still indispensable at the lowest levels. In Indonesia, Malay was initially used by only a small part of the population, though this small part was also the major constituent in the total urbanized and mobile section of the people.

Or there may be several common languages, each functionally inadequate in different ways. To an extent this is the case in the Philippines, where Tagalog is the dominant language of Manila and its surroundings, and Spanish too is taught in the schools, though not more than a few thousands habitually use Spanish. The best example is of course India. English is used for higher education, business and administration, though less than 2 per cent can use it. Hindi is traditionally used in the lower echelons of the armed forces, the majority of the larger market places all over the country, and a large block in the north amounting to about a third of the people, and containing the major share of the common places of religious pilgrimage. Sanskrit is used for ceremonial purposes by all Hindus, with intelligibility by a few thousands. Tamil and Bengali are to some extent used beyond their base territories, the former in the south and the latter in the east, and both have large bodies of speakers outside India, Tamil in Ceylon and Bengali in East Pakistan. If Urdu is considered to be a separate language, which it is only for formal oratory and literature, then we have to say that it is used by about a tenth of the population, almost all of which is fluent in some other Indian language. It too has an important body of users outside the country, namely in West Pakistan. Thus six languages, of which two differ only on formal occasions, qualify as common in different ways. Since eight more languages have official status in regions with considerable political autonomy, a slight change in the point of view would allow us to speak of altogether fourteen common languages (in eleven common scripts).

This kind of a situation may be resolved by several different ways. First, one of the languages may win over the rest. This seems to be the hope in the Philippines. Second, the political unit may be broken up. This happened in the case of the Hapsburg Empire, giving rise to different states, Austria, Hungary, Czechoslovakia, etc. Third, there may be a stabilized bilingualism. This is the case in Finland and Switzerland.

The question of a common language for the sciences is only slightly different. They have become a single tradition, however vaguely and imcompletely defined. But even recently, relatively isolated yet highly creative schools are not uncommon. The logicians of Poland and the linguists of Czechoslovakia between the two world wars made valuable contributions in less widely accessible languages. Undoubtedly, some potential scientific talent is wasted through the dilemma of easier acquisition of science through an individual's first language versus easier acquisition of his results by other scientists through a world language.

There are just two ways of remedying such a situation. One is to step up research in both quantity and quality published in the language concerned, so that scientists elsewhere will soon simply have to learn the language in order to keep up with the field. This is the solution now being attempted in Russia. The other is early and plentiful training in any one of the leading languages of science, so that anyone who has a minimally adequate training in any science is also fluent in such a world language. This solution is slowly coming to prevail in Northern Europe. After a period in which some remarkable work was initially brought out in Dutch and Danish, it is today increasingly less likely that top quality work by Dutch or Danish scientists will be first published in those languages.

The structure of alternatives in the question of scripts, sources of borrowing, etc. is entirely analogous. The alternative that matches most closely the practice of the most accessible of the leading languages of science will always have the entire community of the scientists urging its adoption. This is only limited by the fact that cultivation of science is not the only preoccupation of any social group and that the humanistic sciences exhibit a much less uniform pattern of solidarity. There are also other considerations, such as uniformity between different classes or regions within the nation, neighbors or allies within the commonwealth, generations or traditions within the same history.

A more subtle question arises in relation to the unity of science. Is science a single enterprise at all? The answer is no, if a precisely and enduringly definite body of statements or techniques or points of view is meant. But in a vaguer, more fluid, but nevertheless real sense, the answer may be a differentiated yes. The affirmation is least ambiguous for mathematics, and most for theologies. Yet even for Hindu theology, no one language will be held sufficient today for the purposes of an authoritative scholar. One will need to operate with Sanskrit, Hindi, English and a few other languages. Hence the problems of closure and opening, of the presence or absence of common standards are all there for the languages of even this most marginally scientific of all sciences.

### Questions of policy

The operation of standardization consists basically of two steps, first, the creation of a model for imitation, and second, promotion of this model over rival models.

A model for imitation means a body of discourse capable of attracting the interest and the loyalty of its intended listeners or readers. If we are concerned with the spoken forms, this means the availability of people who habitually use those forms. If we are concerned with the written forms, this means the availability of literature in prose. For either, a model for imitation may be created by unorganized, multiple, obscure decisions, that is, by unplanned decisions. Or it may be done in a relatively planned way. Further, a planned model may be set up on a purely private basis, relying on nothing but persuasion. Or it may be set up on a commissioned basis, relying on more tangible rewards and punishments.

For all practical purposes, the preferred instrument of standardization is officially commissioned publications, with specific conditions as to the choice of the writing system, the source of borrowing, the dialect, etc. The administrative techniques by which literature of the desired kind can be helped into existence are variegated and complex, and need a detailed study.

A few of these administrative techniques may be barely mentioned here. Each has severe limitations in general and for some societies in particular.

1. Approved terminologies. This has its only use in vaguely indicating the general direction of lexical innovation, since particular recommendations can be authoritatively recommended only by those who have assumed responsibility for exhibiting them in consistent and scientifically valid use through an extended body of text.

2. Merit awards. Without attention to specific jobs, persistent and successful work in the desired direction may be acknowledged by public honors or cash grants.

3. Prize contests. A specific job may be chosen and then competition invited for its execution, with materially significant and well-publicized awards for the winners.

4. Nominated authorship. Specific individual writers or translators may be chosen through various means for certain specific pre-chosen jobs. Their work may be supervised and revised by appointed experts.

5. Subsidized self-help. The gap between those who would like to read the book and those who can afford to buy it is never either very small or very large. In the case of books which fulfill the desired conditions it may be bridged by means of a subsidy, the best proven method for which seems to be low-interest loans plus technical advice to private publishers.

6. State enterprise. If a published book is not read, this does not help standardization. It is rather invitingly easy to disguise such waste in State or indeed any non-profit agency enterprise, which tends to be efficient only when publication is mainly for record, as for official or historical or cultural documents.

But the problem is not only to create a model literature in prose but also to promote its acceptance over rival models. Now there is no quick way of making the horse drink the water, if it does not want to. Only the satisfaction of genuinely felt needs can enable a literature to win the allegiance needed. If the information or exposition given is of a kind wanted and is also put in a way that does not deceive or confuse; if for all the insistence on new habits the presentation meets the intended readers more than half way, it may be successful.

Of course, needs may be created or intensified by State efforts. For example, industrialization will certainly intensify the need to have information on all kinds of technology. Or the establishment of a working democracy will create the need for a mass literature on political and economic questions. On the other hand, the decision to withdraw the subsidy for the teaching of Arabic and Persian in Turkey contributed to creation of a need for a more turkified Turkish, just as the decision to penalize any public use of the Arabic-Persian script helped the drive towards general adoption of the Roman script

Need-creating is both relatively useless and relatively harmful without simultaneous model-creating. The stick alone without the carrot may drive the donkey into erratic ways. The argument that once the medium of administration or instruction in India is changed from English to Hindi, satisfactory literature will begin to come out in the language, is not very convincing. Productions that satisfy the best standards are expensive in terms of talent, leisure and money. Productions that satisfy the widest standards of expectation usually do not need subsidies. The essence of the problem is the disparity between the best and the widest standards.

An erratic direction of change may affect the mutual relationships of languages within a multilingual context, with the changes of linguistic habits being either convergent or divergent. Consider for example the question of technical terms in the modern Indian languages. The innovations which are pouring into these languages from English are the same for all and in this respect the change is towards greater convergence. But the technical terms which are being taken from Sanskrit into these languages are rather less convergent. It does not help much if both "sūcanā" and "ghosanā" can be met with in original Sanskrit, if one is selected by Hindi and the other by Bengali for the same English-derived concept "a notice". Even when they are mutually intelligible, and this is not always true, the Sanskrit loanwords in the different languages may be felt as mutually grotesque and barbarous. And to the extent that innovation from non-Sanskrit and non-English sources is preferred, as in Ṭheṭ for Hindi and Urdu, Arabic and Persian for Urdu, Classical Tamil for Tamil etc., the change is towards even greater divergence.

It is of course not necessary for a government to do any kind of planning. It may be just as well done by a church, as was done by Luther and others for German, or by a school of literary artists, as was done by Pramatha Chaudhury and others for Bengali, or by a communion of scientists, as was done by Linnaeus and others for botanical terminology. Even one who prefers to write "phonemic" rather than "phonematic" is doing a bit of prescriptive intervention.

Planning by the State is no guarantee that planning will succeed. One of the clearest examples is in Norway, where the long-ruling Socialists have steadily decried the actual spoken and written standard in favor of more non-Danish forms, but with less and less success. Another example is shown by Papua, where the governor thought that Neo-Melanesian, then known as Melanesian Pidgin, was a disgraceful language and ought to be abolished. The only result of his strenuous efforts has been that the language continues to flourish everywhere except Port Moresby, where it has given way, neither to English nor to any "authentic" (sic) tribal speech, but to Motu, which the governor would consider equally unpedigreed. But perhaps the most celebrated failure has been that of the Irish Republic to reinstate Gaelic.

Success in language planning depends on the already existing network of social communication, that is, on the established channels of commerce in material and intellectual goods. It matters less with whom one is free to communicate than with whom one constantly needs to communicate. Such a need may arise because the other fellow is able to supply a convenience or because association with him gives one self-respect. Prestige derives from both utility and honor. The individuals who participate in relatively more intense communication constitute what Karl Deutsch calls the 'mobilized' part of the population. He has shown how changes in the status of a language no less than in its inventory of forms spread along the lines of social mobilization.

Perhaps a less misleading term for the phenomenon would be 'civilization'. For social mobilization in his sense does not entail anything like official control over newspapers, the radio, the publishing houses etc., or even a bringing together of people in a common task with the roles narrowly and rigidly alloted. What is of concern here are spontaneously formed habits of talking and listening to one another, increasing readiness to explain oneself to or to ask explanation from one another in unrestricted interchange of proposals and comments.

Let us consider some concrete examples. Motu prevailed over other candidates, first because it was based on the speech of the police-boys, who formed the larger part of the total 'mobilized' population, and second because these police-boys were much more frequently and easily accessible to the ordinary Papuans than the English-speaking rulers. Malay prevailed over Javanese and others and was adopted as the Bahasa Indonesia because of its long-standing and wide-ranging circulation by preponderantly Malay-speaking traders, soldiers and priests. Prestige makes people want to adopt a practice, but only access enables them to do it. If different groups of linguistic forms are ordered in a scale of prestige and in one of access, the group which has the highest joint score will have the best chance for general adoption.

In other words, any formal organized action by an acknowledged authority, such as a State or a Church or a learned society or an author, can be successful in its intention to encourage or discourage linguistic habits only if it correlates maximally to informal unorganized action on the part of numerous locally more accessible lesser authorities. Artificial solutions like the IPA script or Esperanto have so little chance of success only because their teachers and texts are so few, indeed initially only one, namely the inventor and his sample composition. In contrast, any native speaker of a natural language functions as some kind of a teacher during the moments of social encounter.

## BIBLIOGRAPHY

Blanc, Haim. "The Growth of Israeli Hebrew." Middle Eastern Affairs (Dec. 1954).

-----. "Hebrew in Israel: Trends and Problems." The Middle East Journal (Autumn 1957).

Deutsch, Karl W. Nationalism and Social Communication. Cambridge, Mass., 1953.

Doke, Clement M. Report on the Unification of the Shona Languages. Hertford, England, 1931.

Fairbanks, Gordon W. "Frequency and Phonemics." Indian Linguistics 17 (1957).

Ferguson, Charles A. "Diglossia." Word 15:2.325-340 (1959).

Ferguson, Charles A. and J. J. Gumperz (eds.) Linguistic Diversity in South Asia. Bloomington, Ind., 1960.

Garvin, Paul L. and M. Mathiot. "The Urbanization of the Guarani Language." In Wallace, Anthony F. C. (ed), International Congress of Anthropological and Ethnological Sciences, 5th., Selected Papers. Philadelphia, 1960.

Gerr, Stanley. Scientific and Technical Japanese. New York, 1944.

Grove, Victor. The Language Bar. New York, 1950.

Gumperz, John J. "Language Problems in the Rural Development of North India." Journal of Asian Studies 16.251-59 (1957)

Harrison, Selig S. The Most Dangerous Decades. New York, 1957.

Haugen, Einar. Bilingualism in the Americas. Alabama, 1956.

Heyd, Uriel. Language Reform in Modern Turkey. Jerusalem, 1954.

Hoijer, Harry A. (ed) Language in Culture. Chicago, 1954.

Joos, Martin. "Semology". Studies in Linguistics 13.53-72 (1958).

Lunt, Horace G. "The Creation of Standard Macedonian." Anthropological Linguistics 1:5 (1959).

Miller, George A. Language and Communication. New York, 1951.

Mills, Harriet C. "Language Reform in China." The Far Eastern Quarterly 15:4.517-540 (1956).

Morag, Shelomo. "Planned and Unplanned Development in Modern Hebrew." Lingua 8:3 (1959).

Prator, Clifford J. Language Teaching in the Philippines. Manila, 1950.

Ray, Punya Sloka. "A Single Script for India." Seminar (July 1960).

-----. "The Definition of a Language." Indian Journal of Philosophy (August 1960).

-----. "The Value of a Language." Lingua 10.220-233 (1961).

-----. "The Formation of Prose." (forthcoming).

-----. "Formal Procedures of Standardization." (forthcoming).

Salzmann, Z. "The Analysis of Lexical Acculturation." International Journal of American Linguistics. 20.137-139 (1954).

Stone, Howard. "Cushioned Loanwords." Word 9:1.1-11 (1953).

UNESCO. The Use of Vernacular Languages in Education. Paris, 1953.

-----. Report on Scientific and Technical Translation. Paris, 1957.

Weinreich, Uriel. Languages in Contact. New York, 1953.

Yngve, Victor H. "A Model and a Hypothesis of Language Structure." Proceedings of the American Philosophical Society 104:5.444-466 (1960).

# SOCIOCULTURAL CHANGE AND COMMUNICATION PROBLEMS

by Janet Roberts

## General Introduction

Rapid social change such as that created by technological innovation disrupts the social structures in those societies in which it occurs and processes of social reorganization must act to bring the various institutional aspects into some functional unity. Prior to change a societal institution ideally manifests a great deal of functional unity and the roles of the members of the group are defined largely by tradition at their birth. However, with change, traditional operations often prove inadequate and special individual talents are required. In general, the social status system shifts from one of role ascription to one of role achievement for effective problem solutions.

Those who have experienced the effects of the industrial revolution are familiar with situations where special occupational talents are required to deal with new problems. But industrialization has only recently begun to affect many areas of the world, and basically stable social groups are encountering situations where the new demands cannot be handled by the existing social order. The most dramatic attempt to cope with this change for many areas in Africa, Asia and Latin America is in the rapid development of educational systems where people can be trained to new occupations. Yet these very areas can be characterized by language complexity and varying degrees of internal, international and scientific/technical communication problems. Here, in the achievement of specific roles, the solution of the language problem becomes one of the first obstacles to be overcome.

The situation in India provides us with an example of language/communication complexity. Three languages are often required for internal communication: the language of birth, the regional language, and Hindi, the official national language. If the individual aspires to higher status in the group he studies a language of technical communication, usually English. For political, business or commercial life it is usually important to learn an Asian language of wider communication, such as Mandarin Chinese or Indonesian, in addition to English. Finally, the university usually requires some Sanskrit.

## Internal, International, and Scientific Communication

It is important to define "internal", "international", and "scientific/technical" communication, with a general statement about the presence or absence of a communication problem in these areas for Africa, Asia and Latin America, before presenting more specific language information for particular countries.[1]

Internal communication is possible for any nation when one language is known by the vast majority of the population. This can occur in any of three ways. First, the country may have a national language which is known by the majority of the people. Second, where several languages are spoken, one of these is dominant and is learned by many of the minority language speakers. Third, if several languages are spoken, and for political or other reasons one of these fails to achieve superordinate status, another language is adopted and learned by most people. Most countries of Asia, Latin America and North Africa have languages which serve the needs of internal communication; however, for Sub-Saharan Africa this is not always the case.

For international communication only a few languages are used extensively. The United Nations officially recognize Mandarin Chinese, English, French, Spanish, and Russian. The fact that the majority of these are European languages can probably be traced to the dominance of Europe over the rest of the world at a time when procedures in trade, business and government were established. A communication problem is created if a people speak a language other than one of the international languages, and of our three areas only Latin America does not have such a problem.

Scientific/technical communication is currently a major problem. UNESCO figures[2] indicate that approximately 71% of all scientific writing is in French, German and English with English accounting for 62% of this output. (One might question the UNESCO figures for scientific materials produced in Russia; the figures may actually be much higher than reported.) At any rate, the scientific and technical material needed to convert basically agricultural states to industrial nations is largely restricted to material in three (or four) languages, and these are languages not known as second languages by enough people in the developing areas to satisfy the need.

The following chart illustrates the previous discussion. Here, the communication problem for Sub-Saharan Africa is the most complicated and the Latin American situation least serious. In the chart a minus sign (-) indicates a problem area.

| | Sub-Saharan Africa | No. Africa and Asia | Latin America |
|---|---|---|---|
| Internal Communication | - | + | + |
| International Communication | - | - | + |
| Scientific/Technical Comm. | - | - | - |

Since English is such an important language of internal, international, and scientific/technical communication, a special statement of the role of English is necessary. English is spoken as a native language, i.e. as the first language learned, by more people than speak any other language natively, with the exception of Chinese. It is the predominating native language of nine nations or territories of more than a million inhabitants: Australia, Canada (67%), England and Wales, Ireland, Jamaica, New Zealand, Northern Ireland, Scotland, United States of America.

For comparative purposes it is useful to list the speakers of those languages spoken natively by the largest numbers of people. The data for this presentation have been taken from the United Nations Demographic Yearbook--1960, Appendix I, except where noted, but no country was included where the population was less than one million. The populations of countries used to arrive at the totals, corrected when necessary, are listed below the table.

| Language | Estimated Number of Native Speakers |
|---|---|
| Chinese | 679,200,000 |
| English | 258,500,000 |
| Hindi-Urdu | 191,000,000 |
| Spanish | 141,600,000 |
| Russian | 124,000,000 |
| Japanese | 92,700,000 |
| German | 83,100,000 |
| Arabic | 76,600,000 |
| Bengali | 75,000,000 |
| Portuguese | 73,300,000 |
| French | 53,600,000 |
| Total | 1,848,600,000 |
| Total World Population | 2,901,834,000 |

*Chinese*. The figure includes the total populations of mainland China (669,000,000) and Taiwan (10,232,000). It should be noted that varieties of Mandarin are probably spoken by some two-thirds of the total and the dialect groups of Wu, Cantonese, Min, and Hakka make up the rest of the figure. This total does not include native speakers of Chinese living in Southeast Asia.

*English*. Australia (10,061,000); Canada (11,686,140) (= 67%); England and Wales (45,504,000); Ireland (2,846,000); Jamaica (1,671,000); New Zealand (2,331,000); Northern Ireland (1,412,000); Scotland (5,241,000); United States (177,700,000). Correction factor for Canada is from UNESCO, *Scientific and Technical Translating and Other Aspects of the Language Problem,* UNESCO, 1958, pp.20-21.

*Hindi-Urdu*. India (185,196,000, = 46% of the population); Pakistan (6,000,000). Hindi and Urdu are sometimes considered to be separate languages: the two differ on formal occasions, e.g. formal oratory and literature; Hindi is written with the Devanagari alphabet, while Urdu is written with a modified form of the Arabic alphabet; for technical, literary, and religious terms, Hindi draws on Sanskrit sources, Urdu on Arabic and Persian sources; Hindi is spoken principally by Hindus, Urdu by Muslims.

*Spanish*. Argentina (20,614,000); Bolivia (1,229,760) (=60%); Chile (7,465,000); Colombia (13,824,000); Costa Rica (1,126,000); Cuba (6,599,000); Dominican Republic (2,894,000); Ecuador (3,585,340) (= 86%); El Salvador (2,016,000) (= 80%); Guatemala (2,191,200) (= 60%); Honduras (1,754,910) (= 93%); Mexico (29,307,520) (= 88%); Nicaragua (1,367,040) (= 96%); Panama (942,080) (= 92%); Paraguay (103,080) (= 6%); Peru (5,682,960) (= 54%); Puerto Rico (1,736,780) (= 74%); Spain (29,894,000); Uruguay (2,700,000); Venezuela (6,512,000). The correction factors are taken from *Statistical Abstract of Latin America - 1961*, Los Angeles; Center of Latin American Studies, University of California, 1961, p. 11.

No accurate per cent for 'Linguistic Composition' is given for El Salvador, Honduras and Peru; the per cent of Indian under 'Ethnic Composition' was used.

*Russian*. The Russian figure is based on the estimate in

**Гиляревский, Р. С. и В. С. Гривнин,** *Опредитель языков мира по письменностям*, **Москва, 1961. Page 11.**

*German*. Austria (7,049,000); East Berlin (1,085,000); East Germany (16,213,000); Federal Republic of Germany (52,785,000); Switzerland (3,772,800) (= 72%); West Berlin (2,211,000). Correction factor for Switzerland is from UNESCO, *op. cit.*, pp. 20-21.

*Arabic*. Algeria (7,400,000) (= 68%); Chad (1,400,000) (= 54%); Egypt (25,300,000); Iraq (5,300,000) (= 76%); Jordan (1,600,000); Lebanon (1,450,000); Libya (1,100,000); Morocco (6,300,000) (= 60%); Saudi Arabia (6,000,000); Sudan (8,000,000) (= 70%); Syria (4,500,000); Tunisia (3,800,000); Yemen (4,400,000).

*Bengali*. India (32,000,000); Pakistan (43,000,000).

*Portuguese*. Brazil (64,216,000); Portugal (9,053,000).

*French*. Belgium (4,096,800) (= 45%); Canada (3,313,980) (= 19%); France (45,097,100); Switzerland (1,048,000) (= 20%). Correction factors are from UNESCO, *op. cit.*, pp. 20-21.

## Problems in Language Communication Research

Research into communication problems may be approached from many different points of view. Basic descriptive data on the languages of a given area according to some typology should be one first step. This is not always easy. First, census data are lacking for many important areas of the world and where data do exist they are often subject to error, do not include information on language speakers, or do not correspond enough to questions of other censuses to permit comparison. Second, certain general information on languages spoken in a given country often is not available from official sources of the country or is completely unknown. Third, information on languages and speakers collected by missionary and other groups has been synthesized to a limited degree only. Fourth, for many languages historical data on the function and number of speakers is lacking altogether. Fifth, the genetic relationship of certain languages of Africa and Latin America is not totally clear. Finally, work on language typology, which is important for descriptive work, has only recently stressed the importance of language function.

In summary, even basic descriptive data on the language situation in a given country or area must be read with caution; it is only recently that enough work has been done to make any such presentation meaningful. The value of any early, general descriptive work must rest not on its reliability but rather on its development of the general problems of communication, its attempt to continue to define significant sociolinguistic concepts, and its function as a base point for future refinement and expansion of both data and concepts.

## Country and Language Data

The purpose of the following section is to illustrate in more detail the communication problem for specific countries of Asia, Africa, and Latin America. For Asia and Africa, where the communication problem is most serious, the data are presented in charts. For Latin America, where the problem is less complex, the information is given in a summary statement.

The data on the charts are set forth under a series of headings: Country and Population. Only countries with a population of over one million are included. Population figures are taken from the United Nations Demographic Yearbook 1960, Appendix I (pp.10-12), except for the figure for Liberia.

*Languages*. In general, only languages with at least 100,000 speakers are listed. This requirement was relaxed for certain important official languages, technical languages, and languages of wider communication. For countries with vast populations, e.g. India, the requirement was raised to one million.

For Africa, the language information is incomplete, but most of the important languages may be assumed to be present. There is great variety in the spelling of the language names, especially for Bantu languages where convention has not yet standardized the treatment of the prefixes *ki-*, *chi-*, etc. Also, a language may have different names in different areas.

In a linguistic sense, *language* and *dialect* have not been differentiated in this survey. Again for Africa, there is some lack of scholarly agreement about the matter.

*Genetic Relationship*. The classification is based principally on Henry A. Gleason, Jr., *An Introduction to Descriptive Linguistics* (Rev. ed., New York, 1961) pp.440-479, supplemented by William E. Welmers' unpublished work sheets on African languages. Genetic relationships for the pidgins and creoles are based on vocabulary.

*Estimated Number of Native Speakers*. In general, no estimates are given for international languages, certain widely known second languages, classical languages, and certain creoles, pidgins, and vernaculars. It should be kept in mind that the figures are estimates only.

*Type and Function*. The material is based on Stewart's typology in this volume (pp.15-25). For convenience, the symbols are repeated here. Language *type* (designated by capital letters): *S*, standard; *C*, classical; *V*, vernacular; *K*, creole; *P*, pidgin. Language *function* (designated by lower case letters): *o*, official; *g*, group; *w*, wider communication; *e*, educational; *l*, literary; *r*, religious; *t*, technical.

*Notes*: The designation "international language" indicates an official language of the U.N.: Chinese, English, French, Russian, Spanish. Diglossia relationships between two languages are symbolized by the notation used in Stewart's typology in this volume (see especially pp. 22-23): language 1 ~ language 2, with the language of higher prestige listed first. The data here are incomplete.

CHART I: ASIA

| Country, Population, Languages | Genetic Relationship (Family-Branch) | Est. No. Speakers (thousands) | Type and Function | Notes |
|---|---|---|---|---|
| AFGHANISTAN (13,150,000) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | | Clr | |
| 2. English | Indo-European (Germanic) | | Sct | Prim. Internat. Lang. |
| 3. Pashto | Indo-European (Iranian) | 6,600 | So | |
| 4. Persian | Indo-European (Iranian) | 3,300 | So | |
| 5. Russian | Indo-European (Slavic) | | St | Internat. Lang. |
| 6. Uzbek | Altaic (Turkic) | 800 | Vg | |
| BURMA (20,457,000) | | | | |
| 1. Burmese | Sino-Tibetan (Tibeto-Burman) | 13,500 | So | |
| 2. English | Indo-European (Germanic) | | Set | Internat. Lang. |
| 3. Indian Languages | | 1,000 | Vg/Sg | |
| 4. Karen Languages | Sino-Tibetan (Tibeto-Burman) | 3,000 | Vg | |
| 5. Shan | Kadai | 1,500 | Vg | |
| CAMBODIA (4,845,000) | | | | |
| 1. Cambodian (Khmer) | Austro-Asiatic | 4,200 | So | |
| 2. Chinese | Sino-Tibeten | 215 | Sg | Internat. Lang. |
| 3. English | Indo-European (Germanic) | | St | Internat. Lang. |
| 4. French | Indo-European (Romance) | | Sot | Prim. Internat. Lang. |
| 5. Pâli | Indo-European (Indic) | | Clr | |
| 6. Vietnamese | Austro-Asiatic | 250 | Sg | |
| CEYLON (9,612,000) | | | | |
| 1. English | Indo-European (Germanic) | | Set | Internat. Lang. |
| 2. Sinhalese | Indo-European (Indic) | 7,000 | So | |
| 3. Tamil (common) | Dravidian | 2,200 | Sco | 4 ~ 3 |
| 4. Tamil (high) | Dravidian | | Cl | 4 ~ 3 |
| CHINA* (Mainland) (669,000,000) | | | | |
| 1. Cantonese | Sino-Tibetan (Chinese) | 46,000 | Sg | Kwang-tung area |
| 2. Hakka | Sino-Tibetan (Chinese) | 17,000 | Vg | South China |
| 3. Lolo | Sino-Tibetan (Chinese) | 3,000 | Vg | |
| 4. Maio | Sino-Tibetan (Chinese) | 3,000 | Vg | South China |
| 5. Mandarin | Sino-Tibetan (Chinese) | 460,000 | So | Internat. Lang. |
| 6. Russian | Indo-European (Slavic) | | St | Internat. Lang. |
| 7. Uighur | Sino-Tibetan | 4,000 | Vg | |
| 8. Wu | Sino-Tibetan (Chinese) | 50,000 | Vg | Yang-tze Delta (Shanghai and Soochow) |
| EGYPT (25,365,000) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | 25,300 | V | 2 ~ 1 |
| 2. Arabic | Afro-Asiatic (Semitic) | | Cor | 2 ~ 1 |
| 3. English | Indo-European (Germanic) | | St | Internat. Lang. |
| 4. French | Indo-European (Romance) | | St | Prim. Internat. Lang. |
| HONG KONG (2,857,000) | | | | |
| 1. Chinese Dialects | Sino-Tibetan | 2,800 | Swe/Vge | |
| 2. English | Indo-European (Germanic) | | Sot | Internat. Lang. |
| INDIA (402,600,000) | | | | |
| 1. Assamese | Indo-European (Indic) | 5,000 | Se | Assam State |
| 2. Bengali | Indo-European (Indic) | 32,000 | Se | W. Bengal & Assam 3 ~ 2 |
| 3. Bengali | Indo-European (Indic) | | Cl | 3 ~ 2 |
| 4. English | Indo-European (Germanic) | | Sot | Internat. Lang. |

---

* It is estimated that there are about 11,700,000 Chinese living abroad and another 8,000,000 in the border regions.

| | | | | |
|---|---|---|---|---|
| INDIA (continued) | | | | |
| 5. Gujerati | Indo-European (Indic) | 20,000 | Se | Gujerati State |
| 6. Hindi-Urdu | Indo-European (Indic) | 185,000 | So | Fig. incl. Bihari, Hindustani, and Punjabi |
| 7. Kannada | Dravidian | 18,000 | Sge | Mysore State |
| 8. Kashmiri | Indo-European (Indic) | 1,500 | Sge | Jammu and Kashmir States |
| 9. Malayalam | Dravidian | 16,000 | Sge | Kerala State |
| 10. Marathi | Indo-European (Indic) | 32,000 | Sge | Bombay Region |
| 11. Oriya | Indo-European (Indic) | 15,000 | Sge | Orissa State |
| 12. Prakrits | Indo-European (Indic) | | Cl | |
| 13. Sanskrit | Indo-European (Indic) | | Clwr | |
| 14. Tamil (common) | Dravidian | 32,000 | Se | 15 ~ 14 Madras State |
| 15. Tamil (high) | Dravidian | | Cl | 15 ~ 14 |
| 16. Telugu | Dravidian | 40,000 | Sge | Andhra Pradesh State |
| INDONESIA (90,300,000) | | | | |
| 1. Balinese | Malayo-Polynesian (Western) | | Cl | 1 ~ 2 |
| 2. Balinese | Malayo-Polynesian (Western) | 2,000 | Vg | 1 ~ 2 |
| 3. Batak | Malayo-Polynesian (Western) | 2,000 | Vg | N. Cent. Reg. Sumatra |
| 4. Bugi | Malayo-Polynesian (Western) | 2,000 | Vg | |
| 5. Chinese Dialects | Sino-Tibetan | 1,600 | Vg | |
| 6. English | Indo-European (Germanic) | | St | Internat. Lang. |
| 7. Indonesian** | Malayo-Polynesian (Western) | 6,700 | So | |
| 8. Javanese | Malayo-Polynesian (Western) | 45,000 | Sg | Cent. and East Java |
| 9. Madurese | Malayo-Polynesian (Western) | 7,000 | Vg | Island of Madura and East Java |
| 10. Minangkabau | Malayo-Polynesian (Western) | 3,000 | Vg | Cent. Sumatra |
| 11. Sundanese | Malayo-Polynesian (Western) | 13,000 | Sg | West. Java |
| IRAN (20,149,000) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | | Clr | |
| 2. Azeri | Altaic | 2,500 | Vg | |
| 3. Kurdish | Indo-European (Iranian) | 500 | Vg | |
| 4. Persian | Indo-European (Iranian) | 14,000 | So | |
| IRAQ (6,952,000) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | | Cor | 1 ~ 2 |
| 2. Arabic | Afro-Asiatic (Semitic) | 5,300 | V | 1 ~ 2 |
| 3. English | Indo-European (Germanic) | | St | Internat. Lang. |
| 4. Kurdish | Indo-European (Iranian) | 1,000 | Ve | |
| ISRAEL (2,061,000) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | | Clr | 1 ~ 2 |
| 2. Arabic | Afro-Asiatic (Semitic) | 200 | Vg | 1 ~ 2 |
| 3. Aramaic | Afro-Asiatic (Semitic) | | Clr | |
| 4. English | Indo-European (Germanic) | | Swt | Prim. Internat. Lang. |
| 5. French | Indo-European (Romance) | | St | Internat. Lang. |
| 6. Hebrew | Afro-Asiatic (Semitic) | 1,800 | So | |
| 7. Judeo-Spanish (Ladino) | Indo-European (Romance) | | V | Speakers incl. under Hebrew |
| 8. Yiddish | Indo-European (Germanic) | | S | Speakers incl. under Hebrew |
| JAPAN (92,740,000) | | | | |
| 1. English | Indo-European (Germanic) | | St | Internat. Lang. |
| 2. Japanese | Japanese | 92,700 | So | 3 ~ 2 |
| 3. Japanese | Japanese | | Clr | 3 ~ 2 |
| 4. Korean | Korean | 600 | Sg | Registered Foreigners |
| JORDAN (1,636,000) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | | Co | 1 ~ 2 |

---

** The figure for Indonesian is expected to remain unreliable for some time. Once a minor language, it was selected through compromise as the official language, and its use in the school system together with its general acceptance by the population should result in a great increase both in native speakers and those who learn it early in life as a second language.

| | | | | |
|---|---|---|---|---|
| JORDAN (continued) | | | | |
| 2. Arabic | Afro-Asiatic (Semitic) | 1,600 | V | 1 ~ 2 |
| 3. English | Indo-European (Germanic) | | St | Internat. Lang. |
| KOREA, NORTH (8,100,000) | | | | |
| 1. Japanese | Japanese | | S | Widely known |
| 2. Korean | Korean | 8,100 | So | |
| 3. Russian | Indo-European (Slavic) | | St | Internat. Lang. |
| KOREA, SOUTH (23,848,000) | | | | |
| 1. English | Indo-European (Germanic) | | St | Internat. Lang. |
| 2. Japanese | Japanese | | S | Widely known |
| 3. Korean | Korean | 23,848 | So | |
| LAOS (1,760,000) | | | | |
| 1. English | Indo-European (Germanic) | | St | Internat. Lang. |
| 2. French | Indo-European (Romance) | | Swet | Prim. Internat. Lang. |
| 3. Lao | Kadai | 1,162 | So | |
| LEBANON (1,550,000) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | | Co | 1 ~ 2 |
| 2. Arabic | Afro-Asiatic (Semitic) | 1,450 | V | 1 ~ 2 |
| 3. Armenian | Indo-European (Armenian) | 100 | Sg | |
| 4. English | Indo-European (Germanic) | | Swt | Internat. Lang. |
| 5. French | Indo-European (Romance) | | Sw | Prim. Internat. Lang. |
| MALAYA (6,698,000) | | | | |
| 1. Chinese Dialects | Sino-Tibetan | 2,473 | Seg/Vge | |
| 2. English | Indo-European (Germanic) | | Sot | Internat. Lang. |
| 3. Indian Lang. (Tamil) | | 804 | Sg/Vg | |
| 4. Malay | Malayo-Polynesian (Western) | 3,349 | So | |
| MONGOLIAN PEOPLE'S REPUBLIC (1,057,000) | | | | |
| 1. Mandarin | Sino-Tibetan (Chinese) | | Sg | Internat. Lang. |
| 2. Mongolian | Altaic (Mongol) | 1,050 | So | |
| 3. Russian | Indo-European (Slavic) | | St | Prim. Internat. Lang. |
| NEPAL (9,044,000) | | | | |
| 1. English | Indo-European (Germanic) | | St | Internat. Lang. |
| 2. Hindi | Indo-European (Indic) | | S | |
| 3. Nepali | Indo-European (Indic) | 4,500 | So | |
| 4. Newari | Sino-Tibetan | | V | |
| PAKISTAN (86,823,000) | | | | |
| 1. Baluchi | Indo-European | 1,000 | V | |
| 2. Bengali | Indo-European (Indic) | 43,000 | S | Regional official<br>3 ~ 2 |
| 3. Bengali | Indo-European (Indic) | | Cl | 3 ~ 2 |
| 4. English | Indo-European (Germanic) | | Sot | Internat. Lang. |
| 5. Kashmiri | Indo-European (Indic) | 1,500 | V | |
| 6. Pashto | Indo-European (Iranian) | 6,400 | V | |
| 7. Punjabi | Indo-European (Indic) | 12,000 | S | |
| 8. Sindhi | Indo-European (Indic) | 5,000 | S | |
| 9. Urdu | Indo-European (Indic) | 20,000 | S | Regional official |
| PHILIPPINES (24,718,000) | | | | |
| 1. Bicol | Malayo-Polynesian (Western) | 2,000 | Vg | |
| 2. Chinese Dialects | Sino-Tibetan | 200 | S/Vg | |
| 3. English | Indo-European (Germanic) | | Sot | Internat. Lang. |
| 4. Ilocano | Malayo-Polynesian (Western) | 2,400 | Vg | |
| 5. Nat'l. Tágalog | Malayo-Polynesian (Western) | 5,000 | So | |
| 6. Pampango | Malayo-Polynesian (Western) | 800 | Vg | |
| 7. Pangasinan | Malayo-Polynesian (Western) | 800 | Vg | |
| 8. Spanish | Indo-European (Romance) | | So | Diminishing use |
| 9. Visayan Group | Malayo-Polynesian (Western) | 10,000 | Vg | |

| | | | | |
|---|---|---|---|---|
| SAUDI ARABIA (6,036,000) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | | Cor | 1 ~ 2 |
| 2. Arabic | Afro-Asiatic (Semitic) | 6,000 | V | 1 ~ 2 |
| 3. English | Indo-European (Germanic) | | St | Internat. Lang. |
| SINGAPORE (1,580,000) | | | | |
| 1. Chinese | Sino-Tibetan (Chinese) | 1,185 | Sge | |
| 2. English | Indo-European (Germanic) | | Sot | Internat. Lang. |
| 3. Malay | Malayo-Polynesian (Western) | 221 | Sge | |
| 4. Tamil (common) | Dravidian | 142 | Sge | 5 ~ 4 |
| 5. Tamil (high) | Dravidian | | Cl | 5 ~ 4 |
| SYRIA (4,539,000) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | | Cor | 1 ~ 2 |
| 2. Arabic | Afro-Asiatic (Semitic) | 4,500 | V | 1 ~ 2 |
| TAIWAN (10,232,000) | | | | |
| 1. Chinese (Mandarin) | Sino-Tibetan (Chinese) | 2,000 | So | Prim. Internat. Lang. |
| 2. Chinese (Amoy, Swatow, Hakka) | Sino-Tibetan (Chinese) | 8,000 | Vg | (original Taiwanese) |
| 3. English | Indo-European (Germanic) | | St | Internat. Lang. |
| 4. Japanese | Japanese | | St | Widely known 2nd Lang. |
| 5. Malay | Malayo-Polynesian (Western) | 100 | Vg | |
| THAILAND (21,881,000) | | | | |
| 1. Chinese Dialects | Sino-Tibetan (Chinese) | 3,000 | Sg | |
| 2. English | Indo-European (Germanic) | | St | Internat. Lang. |
| 3. Malay | Malayo-Polynesian (Western) | 200 | Vg | |
| 4. Thai | Kadai | 18,000 | So | |
| TURKEY (26,881,000) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | | Cr | 1 ~ 2 |
| 2. Arabic | Afro-Asiatic (Semitic) | 350 | Vg | 1 ~ 2 |
| 3. English | Indo-European (Germanic) | | St | Prim. Internat. Lang. |
| 4. French | Indo-European (Romance) | | Sw | Internat. Lang. |
| 5. Kurdish | Indo-European (Iranian) | 1,600 | Vg | Eastern Prov. |
| 6. Turkish | Altaic (Turkic) | 24,600 | So | |
| VIET-NAM, NORTH (15,170,000) | | | | |
| 1. Chinese (Mandarin) | Sino-Tibetan | | St | Internat. Lang. |
| 2. Muong | Sino-Tibetan | 300 | Vg | |
| 3. Thai | Kadai | 500 | Vg | |
| 4. Vietnamese | Austro-Asiatic | 12,900 | So | |
| VIET-NAM, SOUTH (13,790,000) | | | | |
| 1. Chinese | Sino-Tibetan (Chinese) | 126 | S/Vg | No. excludes those with Vietnamese citizenship |
| 2. English | Indo-European (Germanic) | | St | Internat. Lang. |
| 3. French | Indo-European (Romance) | | St | Prim. Internat. Lang. |
| 4. Khmer (Cambodian) | Austro-Asiatic | 300 | Sg | |
| 5. Vietnamese | Austro-Asiatic | 12,400 | So | |
| YEMEN (4,500,000) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | | Cor | 1 ~ 2 |
| 2. Arabic | Afro-Asiatic (Semitic) | 4,400 | V | 1 ~ 2 |

CHART II: AFRICA

| Country, Population, Languages | Genetic Relationship (Family-Branch) | Est. No. Speakers (thousands) | Type and Function | Notes |
|---|---|---|---|---|
| ALGERIA (10,930,000) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | 7,400 | V | 2 ~ 1 |
| 2. Arabic | Afro-Asiatic (Semitic) | | Cor | 2 ~ 1 |
| 3. Berber | Afro-Asiatic (Berber) | 2,500 | Vg | |
| 4. French | Indo-European (Romance) | 1,000 | Sot | Internat. Lang. |
| ANGOLA (4,550,000) | | | | |
| 1. Chokwe (Kioko) | Niger-Congo (Bantu) | 600 | Vg | |
| 2. French | Indo-European (Romance) | | St | Internat. Lang. |
| 3. Kimbundu (Ndongo) | Niger-Congo (Bantu) | 1,000 | Vg | |
| 4. Kongo Dialects | Niger-Congo (Bantu) | 500 | Vg | |
| 5. Portuguese | Indo-European (Romance) | | So | |
| 6. Umbundu | Niger-Congo (Bantu) | 1,700 | Vg | |
| CAMEROUN (4,092,000) (incl. S. British Cameroon) | | | | |
| 1. Bamileke | Niger-Congo (Bantu?) | 500 | Vg | SW |
| 2. Cameroun Pidgin Eng. | Indo-European (Germanic) | | Pw | 4 ~ 2 Coast |
| 3. Duala | Niger-Congo (Bantu) | 300 | Vge | |
| 4. English | Indo-European (Germanic) | | Sot | 4 ~ 2 Internat. Lang. |
| 5. Fang-Bulu | Niger-Congo (W. Atlantic) | 1,700 | Vge | South |
| 6. French | Indo-European (Romance) | | Sot | Internat. Lang. |
| 7. Fulani | Niger-Congo (W. Atlantic) | 1,000 | Vg | Plateau Reg. |
| 8. Hausa | Afro-Asiatic (Chad) | 250 | Sw | |
| 9. Sango-Ngbandi | Niger-Congo | 700 | Vg | |
| CENTRAL AFRICAN REPUBLIC (Ubangi-Shari) (1,185,000) | | | | |
| 1. French | Indo-European (Romance) | | Sot | Internat. Lang. |
| 2. Pidgin Sango | Niger-Congo | | Pw | 3 ~ 2 |
| 3. Sango-Ngbandi | Niger-Congo | 700 | Vgw | 3 ~ 2 |
| 4. Zande | Niger-Congo | 100 | Vg | |
| CHAD (2,600,000 (Tchad) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | 1,400 | Vg | No. Cent. Reg. 2 ~ 1 |
| 2. Arabic | Afro-Asiatic (Semitic) | | Clr | 2 ~ 1 |
| 3. French | Indo-European (Romance) | | Sot | Internat. Lang. |
| 4. Kanuri (Kanembu) | Central Saharan | 330 | Vg | |
| 5. Sara (Sarwa) | Afro-Asiatic (Chad) | 500 | Vg | |
| CONGO (13,821,000) | | | | |
| 1. Banda | Niger-Congo | 150 | Vg | |
| 2. French | Indo-European (Romance) | | Sot | Internat. Lang. |
| 3. Kongo Dialects | Niger-Congo (Bantu) | 800 | Vgwe | SW to Leopoldville |
| 4. Lingala | Niger-Congo (Bantu) | 1,200 | Vgw | Leopoldville North |
| 5. Luba Katanga | Niger-Congo (Bantu) | 800 | Vge | South Cent. |
| 6. Luba Lulua | Niger-Congo (Bantu) | 3,000 | Vgwe | South |
| 7. Mongo Nkundu | Niger-Congo (Bantu) | 500 | Vge | |
| 8. Ruanda | Niger-Congo (Bantu) | 2,400 | Vg | |
| 9. Rundi | Niger-Congo (Bantu) | 1,175 | Vg | |
| 10. Sango-Ngbandi | Niger-Congo | 200 | V | |
| 11. Swahili | Niger-Congo (Bantu) | 2,000 | Swe | South and East 11 ~ 12,13 |
| 12. Swahili | Niger-Congo (Bantu) | | Pw | 11 ~ 12,13 |
| 13. Swahili | Niger-Congo (Bantu) | 2,000 | Vge | South and East 11 ~ 12,13 |
| 14. Zande | Niger-Congo | 650 | Vge | |
| DAHOMEY (2,000,000) | | | | |
| 1. Bariba (Mossi) | Niger-Congo (Gur) | 300 | Vg | NW |
| 2. Ewe-Fon | Niger-Congo (Kwa) | 1,100 | Vg | South |
| 3. French | Indo-European (Romance) | | Sot | Internat. Lang. |
| 4. Fulani | Niger-Congo (W. Atlantic) | 100 | Vg | North |
| 5. Hausa | Afro-Asiatic (Chad) | 115 | Sw | |
| 6. Yoruba | Niger-Congo (Kwa) | 150 | Vg | South |

| | | | | |
|---|---|---|---|---|
| ETHIOPIA (21,800,000) | | | | |
| 1. Amharic | Afro-Asiatic (Semitic) | 7,000 | So | |
| 2. Arabic | Afro-Asiatic (Semitic) | 200 | Vg | 3 ~ 2 |
| 3. Arabic | Afro-Asiatic (Semitic) | | Clr | 3 ~ 2 |
| 4. English | Indo-European (Germanic) | | Sewt | Prim. Internat. Lang. |
| 5. French | Indo-European (Romance) | | St | Internat. Lang. |
| 6. Galla | Afro-Asiatic (Cushitic) | 9,500 | Vg | |
| 7. Gurage | Afro-Asiatic (Semitic) | 350 | Vg | |
| 8. Italian | Indo-European (Romance) | | S | |
| 9. Sidamo | Afro-Asiatic (Semitic) | 1,000 | Vg | |
| 10. Tigre | Afro-Asiatic (Semitic) | 500 | Vg | |
| 11. Tigrinya | Afro-Asiatic (Semitic) | 1,500 | Vg | |
| GHANA (4,911,000) | | | | |
| 1. Akan (Twi, Fante) | Niger-Congo (Kwa) | 2,850 | Vgwe | Coastal Reg. |
| 2. Dagari | Niger-Congo (Gur) | 120 | Vg | |
| 3. English | Indo-European (Germanic) | | Sot | Internat. Lang. |
| 4. Ewe-Fon | Niger-Congo (Kwa) | 150 | Vge | South |
| 5. Ga | Niger-Congo (Kwa) | 250 | Vgwe | Lang. of Capital |
| 6. Moshi-Dagbane | Niger-Congo (Gur) | 180 | Vgwe | North |
| 7. Nzema (Nzima) | Niger-Congo (Kwa) | | Vgwe | West |
| GUINEA (2,727,000) | | | | |
| 1. English | Indo-European (Germanic) | | St | Internat. Lang. |
| 2. French | Indo-European (Romance) | | Sot | Prim. Internat. Lang. |
| 3. Fulani | Niger-Congo (W. Atlantic) | 900 | Vge | Dominant in NW |
| 4. Kissi | Niger-Congo (W. Atlantic) | 250 | Vge | |
| 5. Kpelle (Guerzé) | Niger-Congo (Mande) | 150 | Vg | |
| 6. Malinké | Niger-Congo (Mande) | 600 | Vg | NE |
| 7. Susu | Niger-Congo (Mande) | 312 | Vg | Lang. of Capital and SW |
| IVORY COAST (3,103,000) | | | | |
| 1. Anyi-Baulé (Agni) | Niger-Congo (Kwa) | 975 | Vge | SE |
| 2. Bete (Beti) | Niger-Congo (Kwa) | 250 | Vg | West |
| 3. Dioula (Mandingo Dialect) | Niger-Congo (Mande) | 1,000 | Vge | |
| 4. English | Indo-European (Germanic) | | St | Internat. Lang. |
| 5. French | Indo-Duropean (Romance) | | Sot | Prim. Internat. Lang. |
| 6. Lobi | Niger-Congo (Gur) | 100 | Vg | East |
| 7. Mossi | Niger-Congo (Gur) | | Vg | |
| 8. Senufo | Niger-Congo | 400 | Vg | North |
| KENYA (6,450,000) | | | | |
| 1. English | Indo-European (Germanic) | | Sot | Internat. Lang. |
| 2. Kamba | Niger-Congo (Bantu) | 600 | Vge | |
| 3. Kikuyu | Niger-Congo (Bantu) | 1,300 | Vge | |
| 4. Luo | Chari-Nile (Nilotic) | 800 | Vge | |
| 5. Nandi | Chari-Nile (Nilotic) | 430 | Vge | |
| 6. Swahili | Niger-Congo (Bantu) | (700) | Swe | 6 ~ 7 |
| 7. Swahili | Niger-Congo (Bantu) | (700) | Vg | 6 ~ 7 |
| 8. Teso (Tesoit) | Chari-Nile | 250 | Vge | |
| LIBERIA (1,250,000 - 2,500,000) | | | | |
| 1. English | Indo-European (Germanic) | | Sot | Internat. Lang. 1 ~ 7 |
| 2. Gio | Niger-Congo (Mande) | 100 | Vg | |
| 3. Kissi | Niger-Congo (W. Atlantic) | 300 | Vg | |
| 4. Kpelle | Niger-Congo (Mande) | 500 | Vg | Central and NW |
| 5. Kru-Bassa | Niger-Congo (Kwa) | 600 | Vge | Coastal |
| 6. Mandingo (Maninka) | Niger-Congo (Mande) | 300 | Vg | |
| 7. Pidgin English | Indo-European (Germanic) | | Pw | 1 ~ 7 |
| 8. Vai | Niger-Congo (Mande) | 300 | Vge | SW |
| LIBYA (1,172,000) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | 1,100 | V | |
| 2. Arabic | Afro-Asiatic (Semitic) | | Cor | |
| 3. English | Indo-European (Germanic) | | St | Internat. Lang. |
| 4. Italian | Indo-European (Romance) | | Swe | Widely spoken - older generation. |

MALAGASY (5,239,000) (Madagascar)

| | | | | | |
|---|---|---|---|---|---|
| 1. | French | Indo-European (Romance) | | Sot | Widely spoken;<br>Internat. Lang. |
| 2. | Malagasy | Malayo-Polynesian | 5,000 | So | |

MALI (4,300,000)

| | | | | | |
|---|---|---|---|---|---|
| 1. | Arabic | Afro-Asiatic (Semitic) | 180 | Vgw | 2 ~ 1 North |
| 2. | Arabic | Afro-Asiatic (Semitic) | | Cer | 2 ~ 1 |
| 3. | French | Indo-European (Romance) | | Sot | Internat. Lang. |
| 4. | Fulani | Niger-Congo (W. Atlantic) | 600 | Vg | |
| 5. | Mandingo Dialects (Bambara & Malinké) | Niger-Congo (Mande) | 1,400 | Vg | |
| 6. | Mossi | Niger-Congo (Gur) | 450 | Vg | |
| 7. | Senufo | Niger-Congo | 375 | Vg | |
| 8. | Zarma (Songhai) | Niger-Congo | 250 | Vg | |

MOROCCO (10,550,000)

| | | | | | |
|---|---|---|---|---|---|
| 1. | Arabic | Afro-Asiatic (Semitic) | 6,300 | V | 2 ~ 1 |
| 2. | Arabic | Afro-Asiatic (Semitic) | | Cor | 2 ~ 1 |
| 3. | Berber Group | Afro-Asiatic (Berber) | 4,000 | Vg | |
| 4. | English | Indo-European (Germanic) | | St | Internat. Lang. |
| 5. | French | Indo-European (Romance) | | Set | Prim. Internat. Lang. |
| 6. | Spanish | Indo-European (Romance) | | Se | |

MOZAMBIQUE (6,310,000)

| | | | | | |
|---|---|---|---|---|---|
| 1. | Makua | Niger-Congo | 1,300 | Vg | |
| 2. | Nyungwe | Niger-Congo (Bantu) | 700 | Vg | |
| 3. | Portuguese | Indo-European (Romance) | | So | |
| 4. | Shona Dialects | Niger-Congo (Bantu) | 300 | Vg | |
| 5. | Thonga | Niger-Congo (Bantu) | 1,000 | Vg | |
| 6. | Tshwa | Niger-Congo | 650 | Vg | |

NIGER (2,555,000)

| | | | | | |
|---|---|---|---|---|---|
| 1. | Arabic | Afro-Asiatic (Semitic) | 125 | Vg | 2 ~ 1 |
| 2. | Arabic | Afro-Asiatic (Semitic) | | Cr | 2 ~ 1 |
| 3. | Berber (Tuareg) | Afro-Asiatic (Berber) | 250 | Vg | |
| 4. | French | Indo-European (Romance) | | Sot | Internat. Lang. |
| 5. | Fulani | Niger-Congo (W. Atlantic) | 270 | Vg | |
| 6. | Hausa | Afro-Asiatic (Chad) | 900 | Sw | |
| 7. | Kanuri | Central Saharan | 100 | Vg | |
| 8. | Songhai (Zarma) | Afro-Asiatic (Cushitic) | 200 | Vg | |

NIGERIA (34,458,000)

| | | | | | |
|---|---|---|---|---|---|
| 1. | Bini | Niger-Congo (Kwa) | 500 | Vg | |
| 2. | Efik | Niger-Congo (Central) | 1,300 | Vge | |
| 3. | English | Indo-European (Germanic) | | Sot | Prim. Internat. Lang. |
| 4. | French | Indo-European (Romance) | | Sl | Internat. Lang. |
| 5. | Fulani | Niger-Congo (W. Atlantic) | 3,300 | Vge | North |
| 6. | Gbari | Niger-Congo (Kwa) | 160 | Vg | |
| 7. | Hausa | Afro-Asiatic (Chad) | 12,000 | Swe | Reg. off. Lang.-North |
| 8. | Ibo (Igbo) | Niger-Congo (Kwa) | 4,500 | Swe | Reg. off. Lang.-East |
| 9. | Igbira | Niger-Congo (Kwa) | 150 | Vg | |
| 10. | Ijo | Niger-Congo | 160 | Vge | |
| 11. | Kanuri | Central Saharan | 2,000 | Vge | |
| 12. | Margi | Afro-Asiatic (Chad) | 150 | Vg | |
| 13. | Nupe | Niger-Congo (Kwa) | 326 | Vg | |
| 14. | Sobo | Niger-Congo | 110 | Vg | |
| 15. | Tiv | Niger-Congo (Central) | 700 | Vg | |
| 16. | Yoruba | Niger-Congo (Kwa) | 5,300 | Swe | Reg. off. Lang.-West |

FEDERATION OF RHODESIA AND NYASALAND (Nyasaland) (2,770,000)

| | | | | | |
|---|---|---|---|---|---|
| 1. | English | Indo-European (Germanic) | | Sot | Internat. Lang. |
| 2. | Ngoni | Niger-Congo (Bantu) | 300 | Vg | |
| 3. | Nguru | Niger-Congo | 400 | Vg | |
| 4. | Nyanja | Niger-Congo (Bantu) | 1,000 | Vgwe | Used for off. purposes |
| 5. | Tumbuka | Niger-Congo (Bantu) | 180 | Vge | |
| 6. | Yao | Niger-Congo (Bantu) | 400 | Vge | |

| | | | | |
|---|---|---|---|---|
| RHODESIA, NORTHERN (2,360,000) | | | | |
| 1. Bemba | Niger-Congo (Bantu) | 800 | Vge | Used for off. purposes |
| 2. English | Indo-European (Germanic) | | Sot | Prim. Internat. Lang. |
| 3. French | Indo-European (Romance) | | Slt | Internat. Lang. |
| 4. Ila | Niger-Congo (Bantu) | 250 | Vge | Used for off. purposes |
| 5. Kololo | Niger-Congo | 300 | Vge | Used for off. purposes |
| 6. Nyanja | Niger-Congo (Bantu) | 300 | Vgwe | Used for off. purposes |
| 7. Tonga (Thonga) | Niger-Congo (Bantu) | 300 | Vg | Used for off. purposes |
| RHODESIA, SOUTHERN (3,000,000) | | | | |
| 1. English | Indo-European (Germanic) | | Sot | Internat. Lang. |
| 2. Shona Dialects | Niger-Congo (Bantu) | 1,800 | Vge | Karanga used for off. purposes |
| 3. Tabele (Zulu) | Niger-Congo (Bantu) | 250 | Vge | Used for off. purposes |
| RUANDA-URUNDI (4,780,000) | | | | |
| 1. French | Indo-European (Romance) | | Sot | Internat. Lang. |
| 2. Ruanda | Niger-Congo (Bantu) | 3,500 | Vge | |
| 3. Rundi | Niger-Congo (Bantu) | 600 | Vge | |
| 4. Swahili | Niger-Congo (Bantu) | 500 | Vge | |
| SENEGAL (2,550,000) | | | | |
| 1. Dioula (Mandingo Dialect) | Niger-Congo (Mande) | 300 | Vg | |
| 2. French | Indo-European (Romance) | | Sot | Internat. Lang. |
| 3. Fulani (Poular) | Niger-Congo (W. Atlantic) | 600 | Vg | |
| 4. Serer | Niger-Congo (W. Atlantic) | 400 | Vg | |
| 5. Wolof | Niger-Congo (W. Atlantic) | 850 | Vg | |
| SIERRA LEONE (2,400,000) | | | | |
| 1. Bulom | Niger-Congo (W. Atlantic) | 212 | Vg | |
| 2. English | Indo-European (Germanic) | | Sot | 2 ~ 5<br>Prim. Internat. Lang. |
| 3. French | Indo-European (Romance) | | St | Internat. Lang. |
| 4. Kissi | Niger-Congo (W. Atlantic) | 250 | Vg | |
| 5. Krio | Indo-European (Germanic) | | Kw | Widely known 2nd Lang.<br>2 ~ 5 |
| 6. Limba | Niger-Congo (W. Atlantic) | 175 | Vg | |
| 7. Mende | Niger-Congo (Mande) | 900 | Sgwe | Reg. off. Lang.-South |
| 8. Temne | Niger-Congo (W. Atlantic) | 600 | Sgwe | Reg. off. Lang.-North |
| SOMALIA (1,990,000) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | | Cel | |
| 2. Italian | Indo-European (Romance) | | Swe | |
| 3. Somali | Afro-Asiatic (Cushitic) | 1,800 | Vo | |
| SOUTH AFRICA (14,673,000) | | | | |
| 1. Afrikaans | Indo-European (Germanic) | 1,800 | So | |
| 2. English | Indo-European (Germanic) | 1,200 | Sot | Internat. Lang. |
| 3. N. Sotho | Niger-Congo (Bantu) | 1,265 | Vge | |
| 4. S. Sotho | Niger-Congo (Bantu) | 1,150 | Vge | |
| 5. Tsonga | Niger-Congo (Bantu) | 575 | Vge | |
| 6. Tswana | Niger-Congo (Bantu) | 920 | Vge | |
| 7. Venda | Niger-Congo (Bantu) | 460 | Vge | |
| 8. Xhosa | Niger-Congo (Bantu) | 3,335 | Vge | |
| 9. Zulu, Swazi, Ndebele | Niger-Congo (Bantu) | 3,680 | Vge | |
| SUDAN (11,459,000) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | 8,000 | V | 2 ~ 1 |
| 2. Arabic | Afro-Asiatic (Semitic) | | Cor | 2 ~ 1 |
| 3. Dinka | Chari-Nile (Nilotic) | 600 | Vge | |
| 4. English | Indo-European (Germanic) | | Set | Internat. Lang. |
| 5. Nubian Group | Chari-Nile (Nilotic) | 1,600 | Vg | |
| 6. Nuer | Chari-Nile (Nilotic) | 500 | Vge | |
| 7. Shilluk | Chari-Nile (Nilotic) | 100 | Vge | |
| 8. Acooli-Luo | Chari-Nile (Nilotic) | 500 | Vge | |

| | | | | |
|---|---|---|---|---|
| TANGANYIKA (9,076,000) | | | | |
| 1. Chagga | Niger-Congo (Bantu) | 270 | Vge | |
| 2. English | Indo-European (Germanic) | | Sot | Internat. Lang. |
| 3. Gogo | Niger-Congo (Bantu) | 315 | Vg | |
| 4. Ha | Niger-Congo (Bantu) | 300 | Vg | |
| 5. Iraqw | Afro-Asiatic (Cushitic) | 110 | Vg | |
| 6. Masai | Chari-Nile (Nilotic) | 100 | Vg | |
| 7. Nyamwezi-Sukuma | Niger-Congo (Bantu) | 1,600 | Vge | |
| 8. Swahili | Niger-Congo (Bantu) | (4,500) | Swe | 8 ~ 9, 10 |
| 9. Swahili | Niger-Congo (Bantu) | | Pw | 8 ~ 9, 10 |
| 10. Swahili | Niger-Congo (Bantu) | (4,500) | Vge | 8 ~ 9, 10 |
| TOGO (1,442,000) | | | | |
| 1. Bassari | Niger-Congo (W. Atlantic) | | Vg | |
| 2. Ewe | Niger-Congo (Kwa) | 720,000 | Vg | South |
| 3. French | Indo-European (Romance) | | Sot | |
| 4. Hausa | Afro-Asiatic (Chad) | | Sw | |
| 5. Kabre | Niger-Congo (Gur) | | Vg | |
| 6. Kotokoli | Afro-Asiatic (Chad) | | Vg | Cent. Reg. |
| 7. Moba | Niger-Congo (Gur) | | Vg | |
| TUNISIA (3,935,000) | | | | |
| 1. Arabic | Afro-Asiatic (Semitic) | 3,800 | V | 2 ~ 1 |
| 2. Arabic | Afro-Asiatic (Semitic) | | Cor | 2 ~ 1 |
| 3. French | Indo-European (Romance) | | St | Internat. Lang. |
| UGANDA (6,517,000) | | | | |
| 1. Acooli-Luo | Chari-Nile (Nilotic) | 250 | Vg | |
| 2. English | Indo-European (Germanic) | | Sot | Internat. Lang. |
| 3. Lango | Chari-Nile (Nilotic) | 265 | Vg | |
| 4. Luganda | Niger-Congo (Bantu) | 2,500 | So | |
| 5. Lugbara | Chari-Nile (Central Sudanic) | 150 | Vg | |
| 6. Madi | Chari-Nile (Central Sudanic) | 280 | Vg | |
| 7. Nyoro | Niger-Congo (Bantu) | 700 | Vge | |
| 8. Swahili | Niger-Congo (Bantu) | 215 | Sw | |
| 9. Teso | Chari-Nile | 580 | Vg | |
| UPPER VOLTA (3,537,000) | | | | |
| 1. Bisa | Niger-Congo (Mande) | 100 | Vg | |
| 2. Bobo | Niger-Congo (Gur) | 275 | Vg | West |
| 3. French | Indo-European (Romance) | | Sot | Internat. Lang. |
| 4. Fulani | Niger-Congo (W. Atlantic) | 250 | Vg | |
| 5. Gurma | Niger-Congo (Gur) | 130 | Vg | |
| 6. Lobi | Niger-Congo (Gur) | 200 | Vg | |
| 7. Moré (Mossi) | Niger-Congo (Gur) | 2,000 | Vg | Cent. Reg. |
| 8. Soninke | Niger-Congo | 100 | Vg | |

## LATIN AMERICA

The following material on the language situation in Latin America is presented in summary, since the communication problem here is not as complex as that of Africa and Asia. Only countries with over a million in population were used for the analysis.

*Official Language*. Spanish is the official language of the majority of the countries. Exceptions are Portuguese in Brazil, French in Haiti, English in Jamaica, and English in addition to Spanish in Puerto Rico. The official language is the native language of the vast majority of the population for all countries except in Bolivia, Paraguay, and Haiti (where French Creole is dominant). All except Portuguese and French Creole are languages of formal international communication.

Where language problems exist between Indian tribes and the dominant language groups, the communication problem is eased in some cases by widespread bilingualism. A Spanish/English language problem between North and Latin American people exists. The Portuguese/Spanish language problem is lessened by the near mutual intelligibility of the two languages.

*Medium of Instruction*. The medium of instruction in the schools and the official language of the countries are the same. In Puerto Rico, however, English is not the medium of instruction but is taught as a second language.

English is usually a required course for secondary school children throughout Latin America although in some cases a choice exists between French and English. English is usually favored where it is offered as an elective. In Cuba, Russian and Chinese are now offered to some students as well as French and English.

*Indian Languages*. Quechua, Aymará, and Guaraní are the important Indian language groups of South America, especially Guaraní which has some political importance in Paraguay. In Mexico and Guatemala the following languages have over 100,000 speakers: Cakchiquel, Kekchi, Mam, Mixtec, the Nahuatl group, Otomi, Quiche, Totonac, Yucatec and Zapotec.

NOTES

1. For additional information see Second Language Learning as a Factor in National Development in Asia, Africa, and Latin America, Washington, D. C.: Center for Applied Linguistics, 1961.

2. UNESCO, Scientific and Technical Translating and Other Aspects of the Language Problem, New York: UNESCO Publications Center, 1958, p. 15.

3. For further estimates on number of speakers of languages see C. F. and F. M. Voegelin, "Languages Now Spoken by Over a Million Speakers," Anthropological Linguistics, Bloomington: Archives of Languages of the World, Vol. 3, No. 8, 1961, pp. 13-22. This volume and the following were used for language speakers data and language function material. H. Baumann and D. Estermann, Les peuples et les civilisations de l'Afrique, Paris: Payot, 1957; Center for Applied Linguistics, World Language Survey, unpublished; Encyclopedia Britannica, Chicago: William Benton, Publisher, 1958; Encyclopedia Britannica Book of the Year 1961, Chicago: William Benton, Publisher, 1961; H. A. Gleason, Jr., An Introduction to Descriptive Linguistics, rev. ed., New York: Holt, Rinehart and Winston, 1961, pp. 440-79; Joseph H. Greenberg, Studies in African Linguistic Classification, Branford, Conn.: Compass Publishing Company, 1955; India, a Reference Annual 1961, Delhi: Government of India Press, 1961; International African Institute, Handbook of African Languages, London, Oxford University Press, 1953-9; A. Meillet et Marcel Cohen, eds., Les langues du monde, Paris: Centre National de la Recherche Scientifique, 1952; Mario Pei, One Language for the World, New York: The Devin-Adair Company, 1961; Mario Pei, The World's Chief Languages, New York: S. F. Vanni, 1960; The Statesman's Yearbook...for the Year 1961-62, New York: St. Martin's Press; Elizabeth M. Thompson, Other Lands, Other Peoples: a Country by Country Fact Book, Washington, D. C.: National Education Association, 1961; UNESCO, World Survey of Education III, New York: International Documents Service, 1961; William E. Welmers, "Note on the Classification of African Languages", The Linguistic Reporter, Supplement No. 1, May 1959, Washington, D. C.: Center for Applied Linguistics; William E. Welmers, "A Survey of the Major Languages of Africa," The Linguistic Reporter, Supplement No. 1,

May 1959, Washington, D. C.: Center for Applied Linguistics; William E. Welmers, unpublished work-sheets on African languages; The World Almanac and Book of Facts, 1960, New York: New York World Telegram, 1960; The Worldmark Encyclopedia of the Nations, New York: Worldmark Press, Inc. (Harper and Bros.), 1960.

Most countries with embassies in Washington were contacted for information of varying types.